AD_____

Award Number: DAMD17-97-1-7174

TITLE: Novel and Selective Inactivators for Matrix Metalloproteases Involved in Breast Cancer Metastasis

PRINCIPAL INVESTIGATOR: Shahriar Mobashery, Ph.D.

CONTRACTING ORGANIZATION: Wayne State Unviversity
Detroit, Michigan 48202

REPORT DATE: July 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1999	Annual (13 Jun 98 - 12 Jun 99)

4. TITLE AND SUBTITLE
Novel and Selective Inactivators for Matrix Metalloproteases Involved in Breast Cancer Metastasis

5. FUNDING NUMBERS
DAMD17-97-1-7174

6. AUTHOR(S)
Shahriar Mobashery, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Wayne State University
Detroit, Michigan 48202

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Matrix metalloproteases (MMPs) are endopeptidases with a diverse set of physiological and pathological functions. Two of these proteases, MMP-2 and MMP-9, also known as gelatinases A and B, respectively, have been implicated in breast tumor metastasis. This research endeavors to prepare the first mechanism-based inactivators for these enzymes, molecules which are expected to show high selectivity in inactivation of the targeted enzymes. Toward this goal, several such putative inactivators were designed and synthesized in the Mobashery laboratory in the past year. Initial characterization with the enzyme indicated that one of these molecules showed the desired inactivating trait, but it exhibited relatively low affinities for the gelatinases. Work in on-going in the design of derivatives of higher affinity.

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
41

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

_____ Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

_____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

_____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

_____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

_____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

_____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

PI - Signature 6/9/99 Date

Table of Contents

Introduction

Matrix metalloproteases (MMPs) are zinc-dependent endopeptidases that degrade extracellular matrix components. Two of these proteinases, MMP-2 and MMP-9, also known as gelatinases A and B, respectively, are involved in breast tumor metastasis and the process of neovascularization. Evidence indicates that the activity of gelatinases is critical for breast cancer tumor invasion and angiogenesis. We have undertaken a multidisciplinary research approach for inhibition of the gelatinase-mediated tumor cell invasion and angiogenesis using the first novel synthetic mechanism-based inactivators targeted for gelatinases. These inactivators are expected to show inherent high selectivity for gelatinases, as previously demonstrated by us with similar types of inactivators for other families of metalloproteases. The efficacy and selectivity of the inactivators will be evaluated with purified preparations of recombinant gelatinases A and B (MMP-2, MMP-9), stromelysin-1 (MMP-3) and collagenase-3 (MMP-13). Provisions have been made so that the best inactivator of our design would be evaluated in *in vitro* invasion and angiogenesis assays, as well as in animal models for breast tumor metastasis in collaborations with other groups in the foreseeable future.

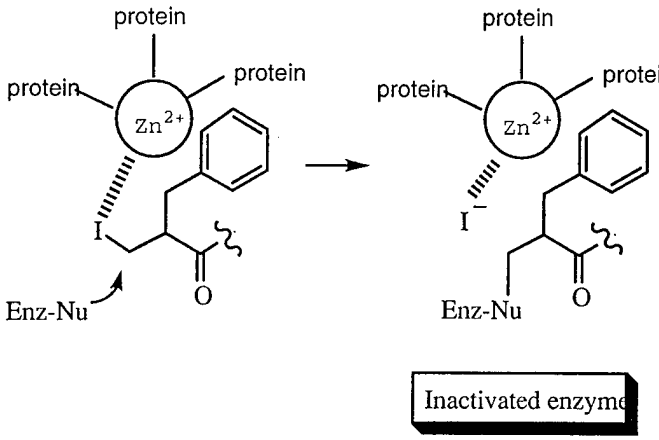
Body

We had reported several inhibitors last year that showed the basic desirable traits of mechanism-based inhibition for gelatinases. The problem rested with their relative poor affinity for the enzyme. We had to increase the affinity, if these strategies were to be successful. We have gone back to the computational models for inhibition of gelatinases and have designed new structural motifs for the inhibitors that we hope would show the mechanism-based inhibition characteristic, as well as high affinity.

The computational models commence with the predicted three-dimensional structure for gelatinases A and B that we reported recently.[1] These models serve as the centerpieces of the design work. We hasten to add that we have made available to the scientific community these models and others at our group Web site (http://sun2.science.wayne.edu/~somgroup/). Our Web site has had over 1000 visitors since late in October of 1999, when we started recording the number of visitors. These visitors often connect to our Web site to download the coordinates for our models. Dr. Dan Ortwine of the Parke-Davis Pharmaceutical Company recently wrote to us to report that they have determined the X-ray structure of the catalytic domain of gelatinase A, and that it matches our model for it. Hence, the models are highly reliable, and they are obviously of use to many research groups.

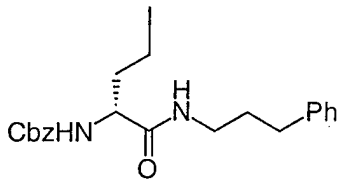
In an aspect of our modeling of these enzymes, we have looked at the diversity of the sequences, and by consequence diversity of functions of MMPs. In the most extensive study of its kind, we just reported a both sequence analysis and modeling effort to understand function of the structures of over 60 MMP.[2] These studies indicated that the accepted dogma for MMPs, that the multiple domains of these proteins assembled to give diversity of function was not valid. The results indicated that assembly of the domains was indeed a very early event in evolution of these enzymes, and diversification of function took place after such initial assembly.[2] Furthermore, we have discovered four distinct motifs for metal binding in MMPs.[3] It is interesting that the three-dimensional structures are strictly conserved, despite the diversity in motifs.[3] These indicate that after assembly of domains, nature went in many different evolutionary tangents to realize distinct functions for the various MMPs. These findings bolster our view that the original proposal by us, that selective inhibition of these enzymes is possible, is indeed valid.

The inhibitors that we set out to make were expected to bind to the gelatinase active site and allow coordination of their halide moiety with the active-site catalytic zinc ion of the enzyme. This interaction would activate the halide for departure. A glutamate is near the zinc ion in the active site, which would be expected to displace the halide. This strategy was implemented successfully for the model zinc-dependent protease, carboxypeptidase A recently,[4,5] and is depicted in the picture below. Modeling with the coordinates of gelatinases has guided us in new directions for design of such inhibitors, as described below.

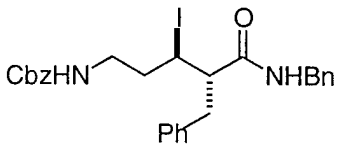


1. Synthesis of iodine containing peptide mimics as inactivators for MMPs

We had to make molecules that were designed for gelatinases specifically. And the interactions had to be of high affinity. Molecular modeling indicates that compounds **1** and **2** should bind to the active site of gelatinases. Furthermore, they would fulfill the coordination chemistry and the predisposition to modification by the glutamate. Compound **1** was synthesized according to Scheme 1.

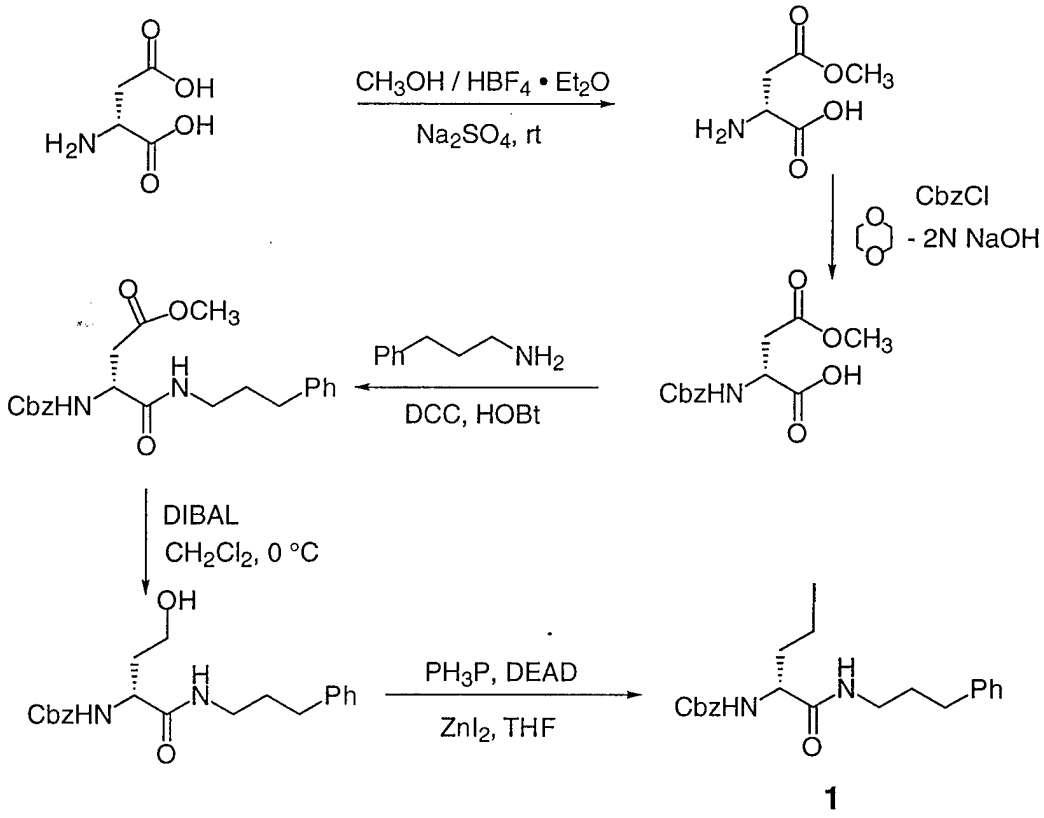


1



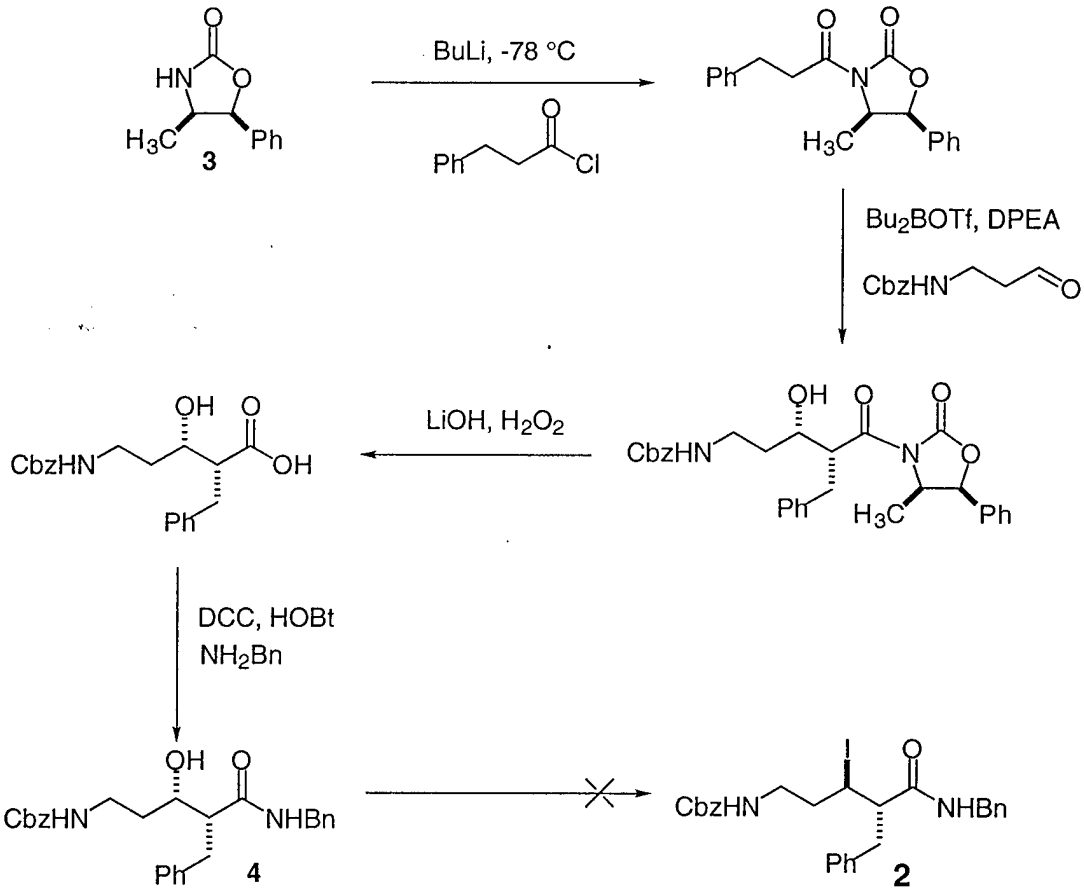
2

Compound **1** inhibited gelatinases according to the profile for mechanism-based inhibition. However, the affinity of the enzymes was not very high for this compound (millimolar range).



Scheme 1

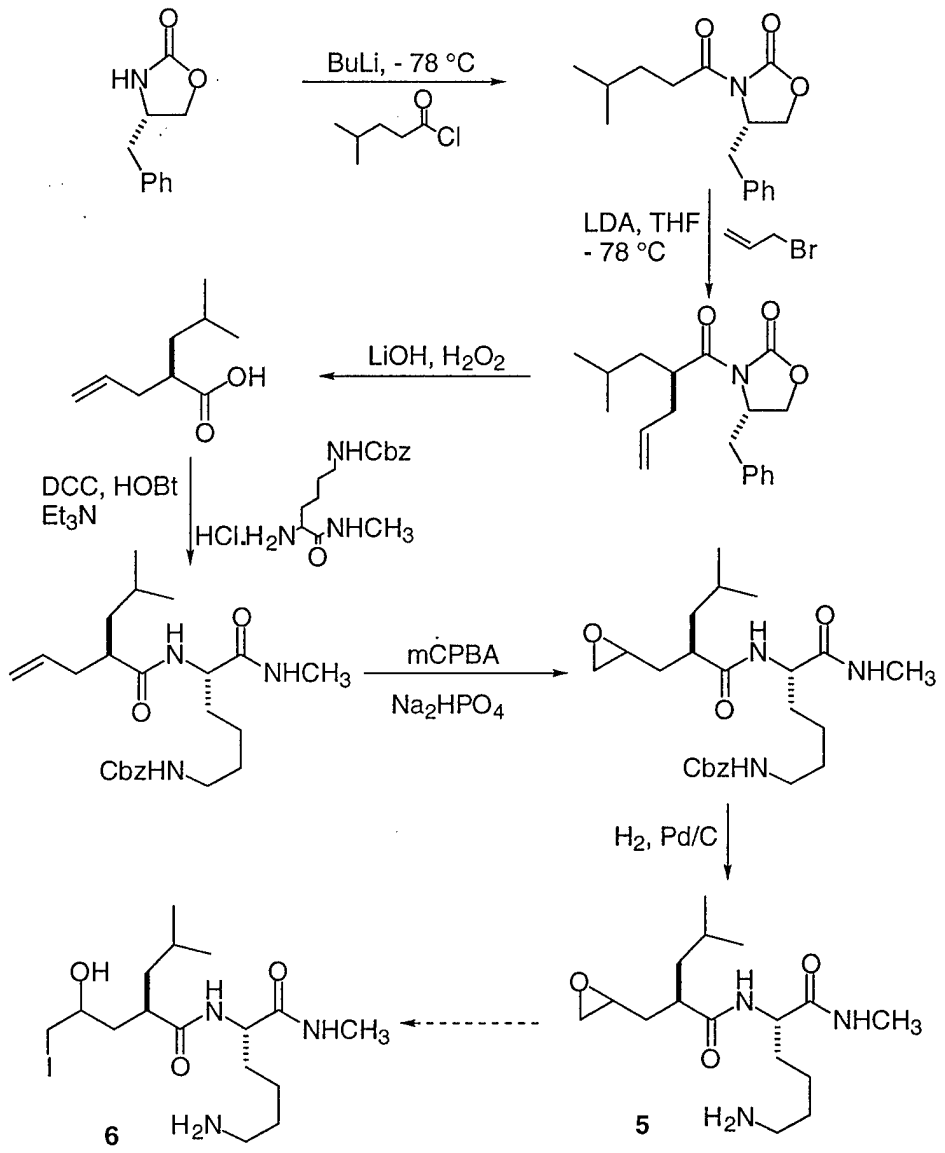
The attempted synthesis of compound **2** (scheme 2) started with the Evans oxazolidinone **3**, and involved a stereospecific aldol reaction. Subsequent cleavage of the auxiliary with lithium hydroxide followed by a DCC coupling reaction with benzylamine led to the precursor **4** in high yield. The conversion of the secondary alcohol of **4** into an iodo group was attempted using numerous methods (over a dozen methds were tried), none of which were successful. The methods attempted included the treatment of the alcohol with :PPh_3, imidazole, I_2; P_2I_4, CS_2; N-methyl-N,N'-dicyclohexylcarbodiimidium iodide; 1-methyl-2-fluoropyridinium p-toluenesulfonate, NaI; 3-ethyl-2-fluorobenzothiazolium tetrafluoroborate, NaI; Tf_2O, pyridine, Bu_4NI; MsCl, pyridine, NaI; PPh_3, DEAD, ZnI_2; PPh_3, DEAD, MeI; TMSCl/NaI, CH_3CN; $SOCl_2$, NaI, CH_3CN; DMF, $(COCl)_2$, NaI.



Scheme 2

2. Synthesis of the epoxide 5 and the hydroxyiodide 6.

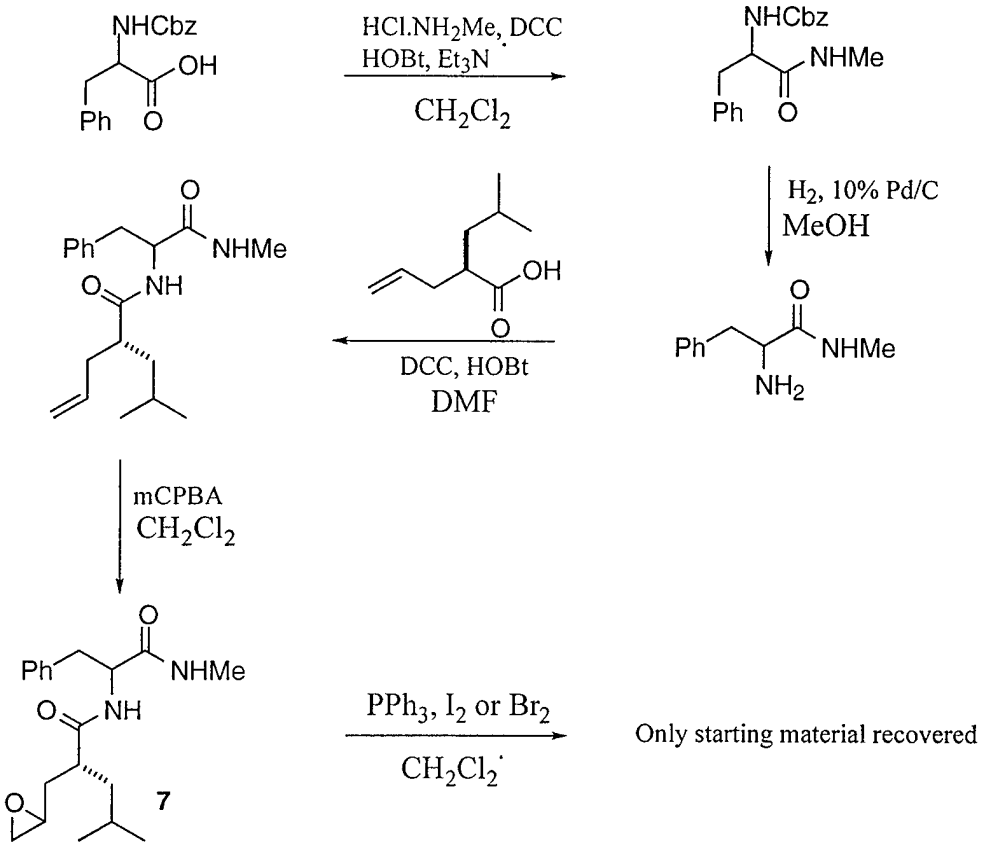
It has been demonstrated that an epoxide group can be a potent and effective candidate for inactivation of carboxypeptidase A, a model zinc protease.[6] The principle is the same as the halide strategy. The epoxide oxygen would coordinate to the active-site zinc, and it is displaced by the incoming nucleophile. For targeting of gelatinase, epoxide of **5** would coordinate to the active-site zinc ion, thus increasing the reactivity of the oxirane towards attack by the glutamate. Subsequent attack of the activated oxirane by the active-site glutamic acid would result in the covalent bonding of **5** to the enzyme, resulting in the loss of enzyme activity. The reason that **6** was chosen as a target molecule was because we expected that both the hydroxyl and iodo groups would coordinate to the active-site zinc ion. This chelation may increase the affinity for the site. Compound **5** was prepared (scheme 3) but all attempts to date to convert the epoxide into the hydroxy iodo **6** failed. Furthermore, enzymatic testing revealed that compound **5** did not inactivate the enzyme.



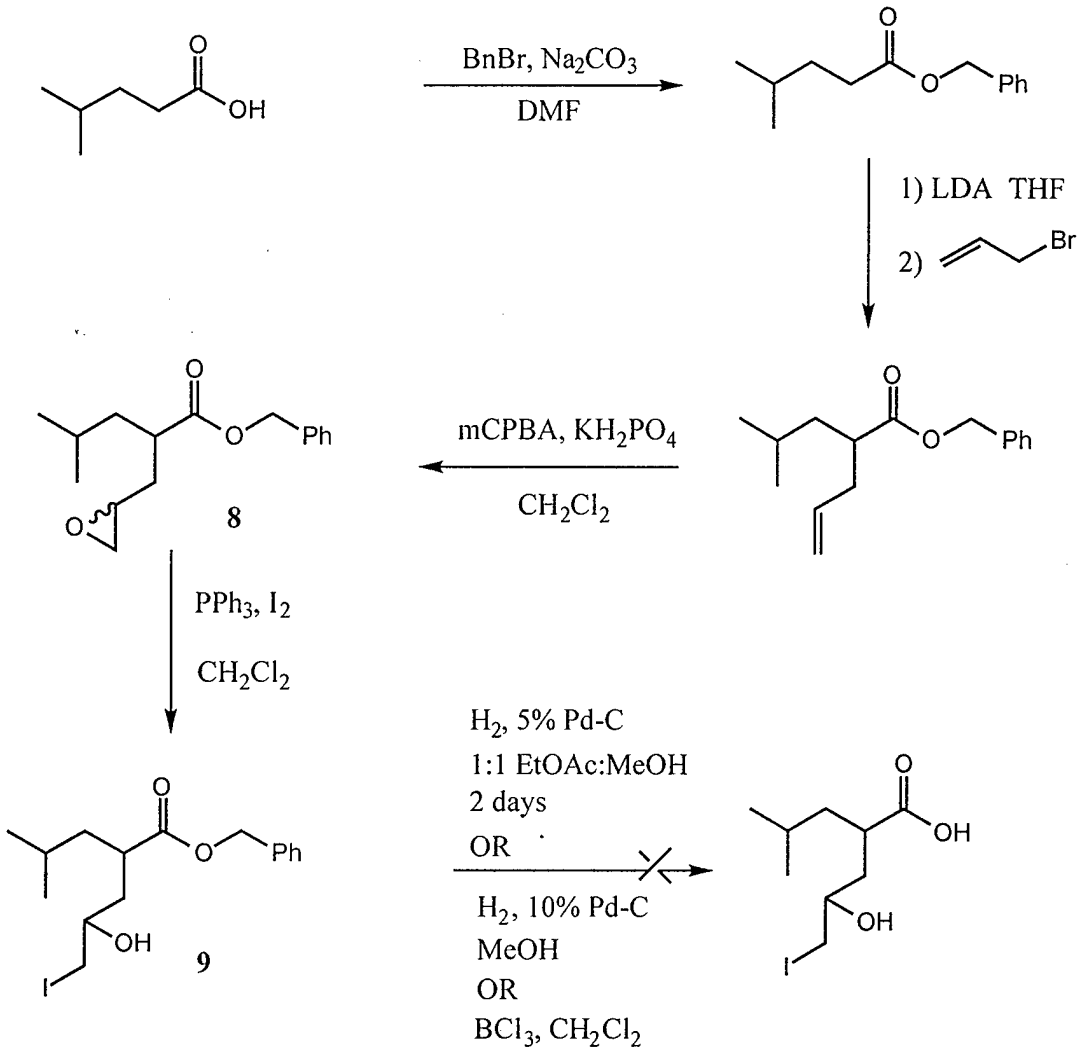
Scheme 3

Similarly, attempts where made to synthesize the peptide based epoxide **7** (scheme 4). At first it seemed as though the reaction to form the epoxide **7** had been successful, but later attempts to assign the ^1H NMR spectrum for this product proved impossible. The ^1H NMR spectrum did appear to show a 1:1 mixture of diastereoisomers, as would be expected for a reaction of this type, but not all the chemical shift values agree with the proposed structure. The reaction to convert the desired epoxide **7** into a halohydrin was

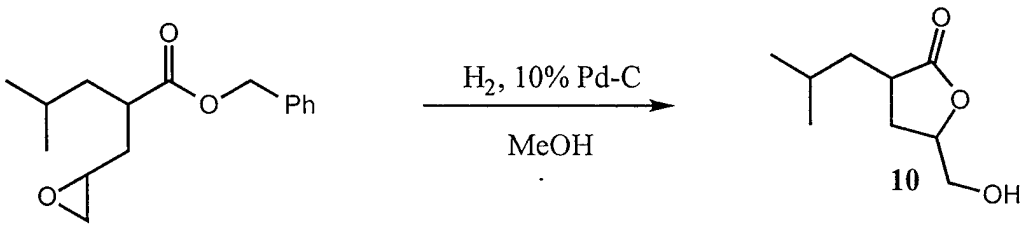
attempted before detailed investigation of the ^1H NMR spectrum of **7** was carried out. Unable to prepare either the epoxide or an iodohydrin via this route, an alternative synthesis was sought (scheme 5). The idea behind this route was to prepare the epoxide or the iodohydrin-containing fragment as an individual compound and then attach it to the peptide component using a DCC coupling reaction. Although we successfully prepared both the desired epoxide **8** and iodohydrin **9**, all attempts to remove the benzyl group failed. When the epoxide was treated with hydrogen in the presence of 10% Pd-C, the lactone **10** was formed (Scheme 6), and when attempts were made to remove the benzyl group from the iodohydrin, a complex mixture of compounds was obtained.



Scheme 4



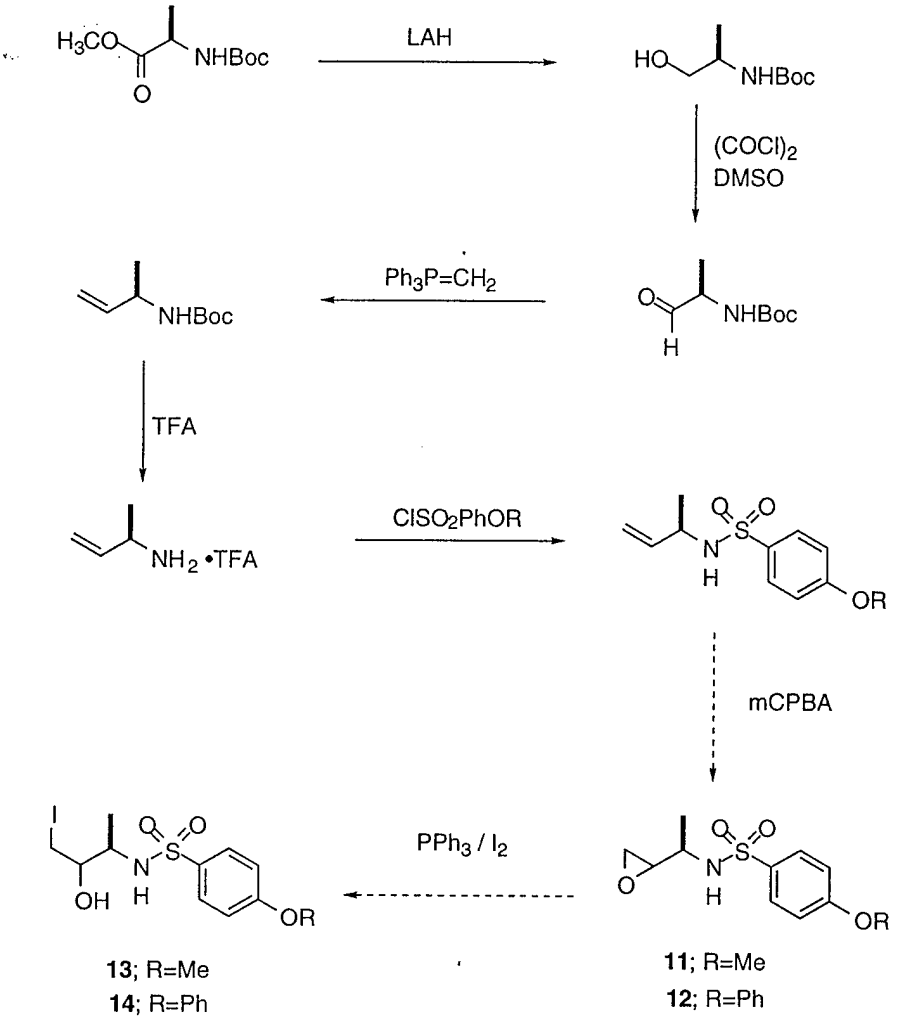
Scheme 5



Scheme 6

3. Sysnthesis of nonpeptide epoxide and hydroxyiodide compounds as Gelatinase inactivators.

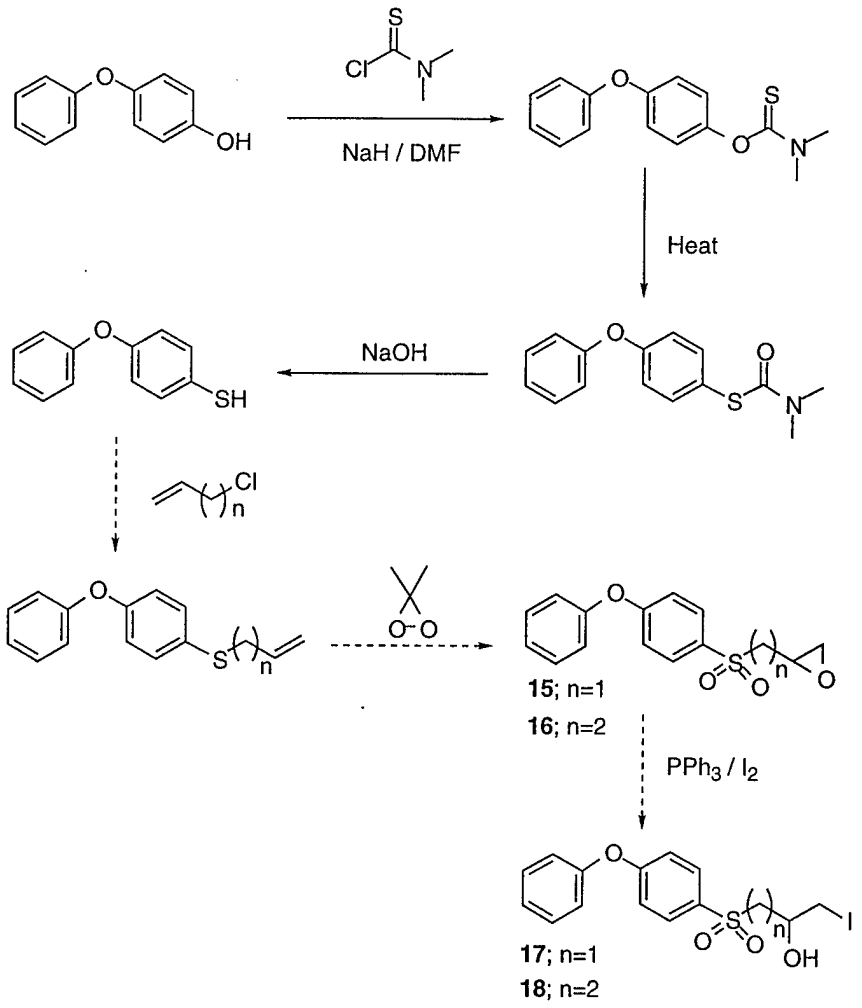
It was found that the peptide mimic epoxide **5** was unstable during purification. It is for this reason that nonpeptide epoxide inactivators (**11**,**12**,**15**,**16**) and hydroxyiodide (**13**,**14**,**17**,**18**) were designed and the syntheses are currently under investigation.



13; R=Me
14; R=Ph

11; R=Me
12; R=Ph

Scheme 7



Scheme 8

Key Research Accomplishments

* Synthesized a number of derivatives leading to proposed target molecules.
* Two of the potential gelatinase inhibitors were successfully synthesized. One showed mechanism-based inhibition of the enzyme, though affinity was low.
* New target molecules were designed using computational modeling.

Reportable Outcomes

* Two papers have resulted from our modeling analyses. They are provided as appendices.
* The database for sequence analyses for MMPs, and the coordinates for their models, are provided to the public in our group Web site.

Conclusions

We have synthesized several molecules, in multiple steps each, as potential inactivators of gelatinases. These molecules establish some of the required elements in binding to the active site, and one show time-dependence for loss of activity, which is

required for the inactivators of our design. Our recent results point us in the direction of improving our design paradigms, which should lead to molecules of high potency for inhibition of gelatinases.

References

(1) Massova, I. ; Fridman, R.; Mobashery, S. Structural Insights into the Catalytic Domains of Human Matrix Metalloprotease-2 and Human Matrix Metalloprotease-9: Implications for Substrate Specificities, *J. Mol. Mod.* **1997**, *3*, 17.

(2) Massova, I.; Kotra, L. P.; Fridman, R.; Mobashery, S. Matrix Metalloproteases: Structures, Evolution and Diversification, *FASEB J.* **1998**, *12*, 1075.

(3) Massova, I.; Kotra, L. P.; Mobashery, S. Structural Insight into the Binding Motifs for Calcium Ion and the Non-Catalytic Zinc in Matrix Metalloproteases, *Bioorganic Med. Chem. Lett.* **1998**, *8*, 853.

(4) Tanaka,Y.; Grapsas, I.; Dakoji, S.; Cho, Y.J.; Mobashery, S. Conscripting the Active-Site Zinc Ion in Carboxypeptidase A in Inactivation Chemistry by a New Type of Irreversible Enzyme Inactivator, *J. Am. Chem. Soc.* **1994**, *116*, 7475.

(5) Massova, I.; Martin, P.; de Mel, S.; Tanaka, Y.; Edwards, B.; Mobashery, S. Crystallographic and Computational Insight on the Mechanism of Zinc-Ion-Dependent Inactivation of Carboxypeptidase A by 2-Benzyl-3-Iodopropanoate, *J. Am. Chem. Soc.* **1996**, *118* , 12479.

(6) Lee, S. S.; Li, Z. H.; Lee, D. H.; Kim, D. H. (2*R*,3*S*)- and (2*S*,3*R*)-2-Benzyl-3,4-Epoxybutanoic Acid as Highly Efficient and Fast Acting Pseudo-Mechanism-Based Inactivator for Carboxypeptidase A. Design, Asymmetric Synthesis, and Inhibitory Kinetics, *J. Chem. Soc., Perkin Trans.* **1995**, *1*, 2877.

Matrix metalloproteinases: structures, evolution, and diversification

IRINA MASSOVA,*[,1] LAKSHMI P. KOTRA,* RAFAEL FRIDMAN,[†] AND SHAHRIAR MOBASHERY*[,2]

*Department of Chemistry, and [†]Department of Pathology and Karmanos Cancer Institute, Wayne State University, Detroit, Michigan 48202–3489, USA

ABSTRACT **A comprehensive sequence alignment of 64 members of the family of matrix metalloproteinases (MMPs) for the entire sequences, and subsequently the catalytic and the hemopexin-like domains, have been performed. The 64 MMPs were selected from plants, invertebrates, and vertebrates. The analyses disclosed that as many as 23 distinct subfamilies of these proteins are known to exist. Information from the sequence alignments was correlated with structures, both crystallographic as well as computational, of the catalytic domains for the 23 representative members of the MMP family. A survey of the metal binding sites and two loops containing variable sequences of amino acids, which are important for substrate interactions, are discussed. The collective data support the proposal that the assembly of the domains into multidomain enzymes was likely to be an early evolutionary event. This was followed by diversification, perhaps in parallel among the MMPs, in a subsequent evolutionary time scale. Analysis indicates that a retrograde structure simplification may have accounted for the evolution of MMPs with simple domain constituents, such as matrilysin, from the larger and more elaborate enzymes.—Massova, I., Kotra, L. P., Fridman, R., Mobashery, S. Matrix metalloproteinases: structures, evolution, and diversification. *FASEB J.* 12, 1075–1095 (1998)**

Key Words: extracellular matrix · MMP · hemopexin · tissue inhibitor of matrix metalloproteinase

BACKGROUND

The interactions of cells with the extracellular matrix (ECM)[3] are critical for the normal development and function of the organism. Modulation of cell–matrix interactions occurs through the action of unique proteolytic systems responsible for hydrolysis of a variety of ECM components. By regulating the integrity and composition of the ECM structure, these enzyme systems play a pivotal role in the control of signals elicited by matrix molecules, which regulate cell proliferation, differentiation, and cell death. The turnover and remodeling of ECM must be highly regulated since uncontrolled proteolysis contributes to abnormal development and to the generation of many pathological conditions characterized by either excessive degradation or a lack of degradation of ECM components. Matrix metalloproteinases (MMPs) are a major group of enzymes that regulate cell–matrix composition. The MMPs are zinc-dependent endopeptidases known for their ability to cleave one or several ECM constituents, as well as nonmatrix proteins. They comprise a large family of proteases that share common structural and functional elements and are products of different genes. Ample evidence exists on the role of MMPs in normal and pathological processes, including embryogenesis, wound healing, inflammation, arthritis, and cancer. The association of MMPs with cancer metastasis has raised considerable interest because they represent an attractive target for development of novel antimetastatic drugs aimed at inhibiting MMP activity. Therefore, understanding the structure and function of these key enzymes has significant implications for cancer therapy (1–5).

Most members of the MMP family are organized into three basic, distinctive, and well-conserved domains based on structural considerations: an amino-terminal propeptide; a catalytic domain; and a hemopexin-like domain at the carboxy-terminal (**Fig. 1**). The propeptide consists of approximately 80–90 amino acids containing a cysteine residue, which interacts with the catalytic zinc atom via its side chain thiol group. A highly conserved sequence (. . .PRCGXPD. . .) is present in the propeptide. Removal of the propeptide by proteolysis results in zymogen activation, as all members of the MMP family are produced in a latent form. The catalytic domain contains two zinc ions and at least one calcium

[1] Current address: University of California at San Francisco, 513, Parnassus, S926, San Francisco, CA 94143–0446, USA.

[2] Correspondence: Shahriar Mobashery, Department of Chemistry, Wayne State University, Detroit, MI 48202–3489, USA. E-mail: som@mobashery.chem.wayne.edu

[3] Abbreviations: ECM, extracellular matrix; MMPs, matrix metalloproteinases; TIMPs, tissue inhibitors of metalloproteinases; MT-MMPs, membrane-type MMPs.

0892-6638/98/0012-1075/$02.25 © FASEB

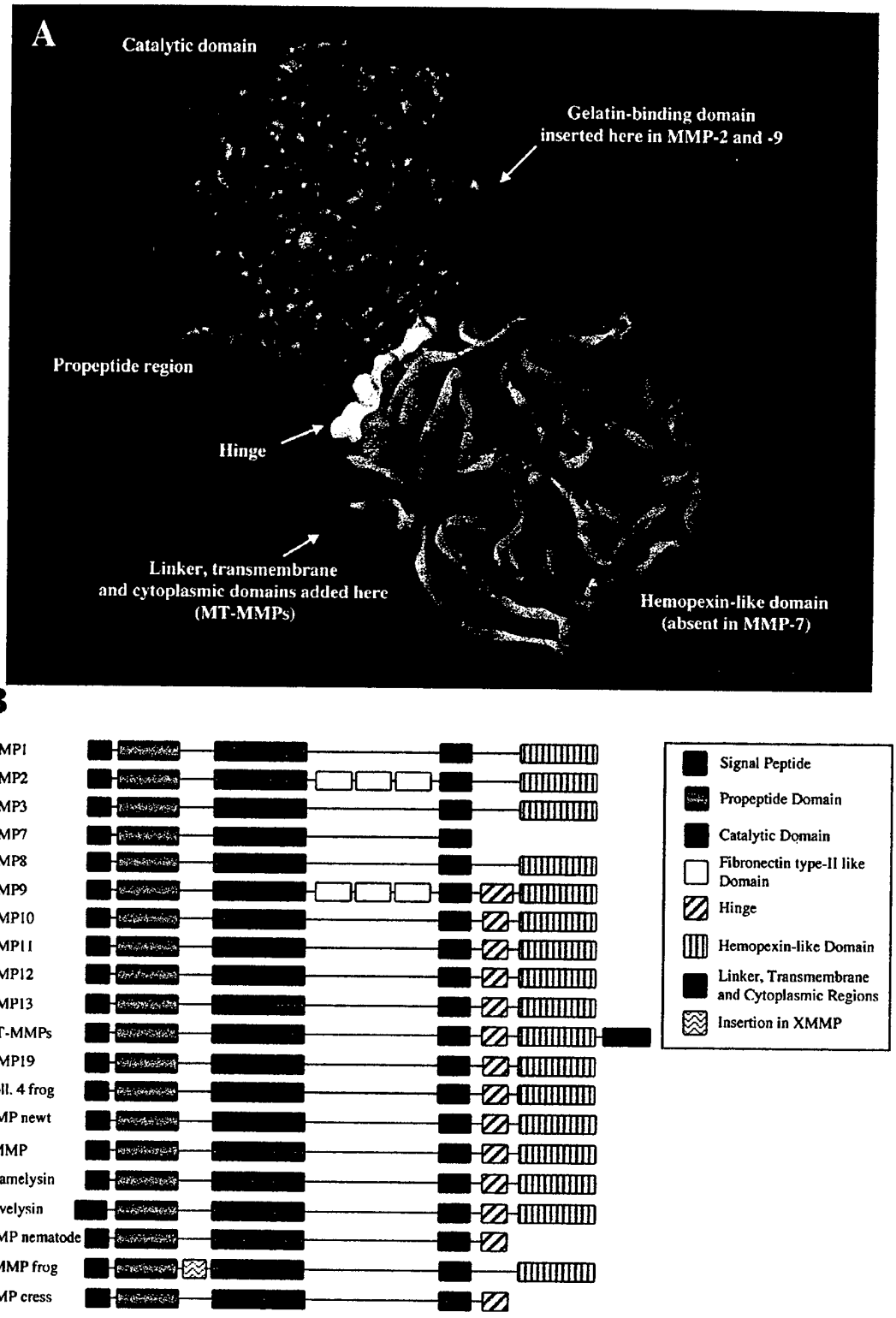


Figure 1. *A)* Basic domain structures of MMPs. The images for the propeptide region and the catalytic and homopexin-like domains shown here are from crystallographic sources. The propeptide region is taken from the X-ray structure for stomelysin (PDB code: 1slm) and the remaining portions of the structure are taken from the X-ray structure of the full-length collagenase (PDB code: 1fbl). Catalytic zinc is shown as an orange sphere; calcium ions in the catalytic domain and the hemopexin-like domain are shown in cyan. The propeptide region is shown by the green ribbon, catalytic domain as a surface in pink, hinge region as a surface in white, and the hemopexin-like domain is represented by the ribbon drawing in yellow. *B)* Schematics of the domain structures of the 23 representative MMPs. Catalytic domain (represented by green) has an insertion of gelatin binding domain in MMP-2 and MMP-9. In all other MMPs, the catalytic domain is a continuous entity.

ion coordinated to various residues. One of the two zinc ions is present in the active site and is involved in the catalytic processes of the MMPs. The second zinc ion (also known as structural zinc) and the calcium ion are present in the catalytic domain approximately 12 Å away from the catalytic zinc. The catalytic zinc ion is essential for the proteolytic activity of MMPs; the three histidine residues that coordinate with the catalytic zinc are conserved among all the MMPs. Little is known about the roles of the second zinc ion and the calcium ion within the catalytic domain, but the MMPs are shown to possess high affinities for structural zinc and calcium ions (6, 7). The hemopexin-like domain of MMPs is highly conserved and shows sequence similarity to the plasma protein, hemopexin. The hemopexin-like domain has been shown to play a functional role in substrate binding and/or in interactions with the tissue inhibitors of metalloproteinases (TIMPs), a family of specific MMP protein inhibitors (8, 9). In addition to these basic domains, the family of MMPs evolved into different subgroups by incorporating and/or deleting structural and functional domains. For example, MMP-2 and MMP-9, also known as gelatinases, incorporated the three repeats homologous to the type-II module of fibronectin into the catalytic domain that has been shown to be involved in binding to denatured collagen or gelatin (10). This domain, known as the gelatin binding domain or fibronectin type-II-like domain, is unique to the gelatinases, and so these enzymes are regarded as a separate subgroup among members of the MMP family. Incorporation of a hydrophobic stretch of approximately 25 amino acids, representing a putative transmembrane domain at the carboxy terminus and recognition motif (RXKR) for furin-like convertases at the end of the propeptide domain, is a characteristic of the membrane-type MMPs (MT-MMPs) (11, 12) except MT4-MMP (vide infra). MMP-11 also contains this furin recognition motif and, similar to the MT-MMPs, it is processed into the active form intracellularly (13). Additional insertion to the three basic MMP domains also includes a proline-rich 54 amino acid insertion in MMP-9 with sequence similarity to the α_2 chain of collagen V (14). Finally, MMP-7 lacks the hemopexin-like domain and represents the smallest member of the MMP family.

The catalytic activity of the MMPs is regulated at multiple levels including transcription, secretion, activation, and inhibition. The last is accomplished by members of the TIMP family, which presently includes four proteins: TIMP-1, TIMP-2, TIMP-3, and TIMP-4 (8, 15). Binding of the TIMPs to the catalytic domain results in efficient inhibition of enzymatic activity of MMPs. In the case of gelatinases, the TIMPs have been shown to bind to the zymogen forms of the enzymes. This interaction has been suggested to provide an extra level of regulation by potentially preventing activation (15, 16). However, it has recently been shown that TIMP-2 forms a trimolecular complex on the surface of the cell with MT1-MMP and proMMP-2, and regulates the formation and levels of concentration of mature MMP-2 (17). The crystal structure of the catalytic domain of MMP-3 in complex with TIMP-1 has been solved and shows that Cys^1 of the inhibitor interacts with the catalytic zinc ion of the MMP through the α-amino and its carbonyl group, whereas the Thr^2 side chain extends into the S_1' specificity pocket of the enzyme (9). A critical step in the control of MMP activity is regulated by the generation of active enzyme species with proteolytic activity. The process of activation involves sequential cleavage of the propeptide, which disrupts coordination between the cysteine thiol in the propeptide region with the zinc atom in the catalytic domain. This process is proteolytic and may involve other MMPs acting in a cascade of zymogen activation.

MMPs belong to the superfamily of zinc-peptidases, and the evolutionary relationship of this superfamily has been reviewed (18, 19). Sang and Douglas (20) analyzed 30 MMP sequences from various sources by multiple-sequence alignment, but their study was limited to analysis of the primary sequences. A total of 66 MMPs have been sequenced to date, of which 17 are from humans, including the recently discovered human enamelysin (MMP-20) and a functional enzyme encoded by the mmp20 gene (GenBank accession number AJ003147, maps to chromosome 16) in the familial Mediterranean fever gene in humans (21, 22). These human enzymes have counterparts in other vertebrates. MMPs have even been identified in invertebrates (23–25) and three have recently been sequenced from plant sources (mouse ear cress MMPs, **Table 1**). MMPs are probably more ancient than is currently realized. Their origin might actually be traceable back to bacteria in that a certain amino acid sequence for *Bacteroides fragilis* metalloproteinase toxin-2 (GenBank accession number U90931) has 59% sequence identity to the continuous 27 amino acid stretch in human MMP-1, which includes the catalytic zinc binding domain and the 'methionine turn' (this is a strictly conserved region with a methionine in the catalytic domain of MMPs responsible for the structural integrity of the zinc binding site). A salient question is why such multiplicity of these enzymes is seen in nature. We wish to add to this question an additional inquiry: What makes them different, and how does the difference in amino acid sequences give rise to structural elements that, in turn, would render a given MMP a distinct enzyme? We have compared amino acid sequences from 64 MMPs from various organisms, vertebrates, invertebrates, and plants to address these questions. Using the available crystal structures for four MMPs (26–29), we have modeled the 3-dimensional structures of several representative members of the remaining MMPs to gain insight into

TABLE 1. *Sources of various MMPs used in multiple-sequence alignment*

Gene, organism	Common name	Accession #	Database
MMP-2 (gelatinase A)			
Mmp2, Mus musculus	Mouse	P33434	Swiss-Prot
Mmp2, Rattus norvegicus	Rat	P33436	Swiss-Prot
MMP2, Homo sapiens	Human	P08253	Swiss-Prot
Mmp2, Oryctolagus cuniculus	Rabbit	P50757	Swiss-Prot
Mmp2, Gallus gallus	Chicken	Q90611	Swiss-Prot
MMP-9 (gelatinase B)			
Mmp9 or Clg4b, Mus musculus	Mouse	P41245	Swiss-Prot
Mmp9, Rattus norvegicus	Rat	P50282	Swiss-Prot
MMP9, Homo sapiens	Human	P14780	Swiss-Prot
Mmp9, Oryctolagus cuniculus	Rabbit	P41246	Swiss-Prot
Mmp9, Bos taurus	Bovin	P52176	Swiss-Prot
Mmp9, Synops pyrrhogaster	Newt	Q98856	TREMBL
MMP-12 (macrophage metalloelastase)			
Rattus norvegicus	Rat	X98517	GenBank
Oryctolagus cuniculus	Rabbit	U88652	GenBank
MMP12 or *HME, Homo sapiens*	Human	P39900	Swiss-Prot
Mmp12 or *Mme, Mus musculus*	Mouse	P34960	Swiss-Prot
MMP-13 (collagenase-3)			
Mmp13, Mus musculus	Mouse	P33435	Swiss-Prot
Mmp13, Rattus norvegicus	Rat	P23097	Swiss-Prot
MMP13, Homo sapiens	Human	P45452	Swiss-Prot
gene A, Xenopus laevis	Frog	U41824	GenBank
Mmp13, Xenopus laevis	Frog	Q10835	Swiss-Prot
Cynops pyrrhogaster	Newt	D82055	GenBank
Collagenase-4			
Xenopus laevis	Frog	L76275	GenBank
MMP-3 (stromelysin-1)			
Mmp3, Mus musculus	Mouse	P28862	Swiss-Prot
Mmp3, Rattus norvegicus	Rat	P03957	Swiss-Prot
MMP3 or *STMY1, Homo sapiens*	Human	P08254	Swiss-Prot
Mmp3, Oryctolagus cuniculus	Rabbit	P28863	Swiss-Prot
Equus caballus	Horse	U62529	GenBank
MMP-10 (stromelysin-2)			
MMP10 or *STMY2, Homo sapiens*	Human	P09238	Swiss-Prot
Mmp10, Rattus norvegicus	Rat	P07152	Swiss-Prot
MMP-1 (interstitial collagenase)			
Mmp1 or *Clg, Bos taurus*	Bovin	P28053	Swiss-Prot
Mmp1, Sus scrofa	Pig	P21692	Swiss-Prot
MMP1 or *CLG, Homo sapiens*	Human	P03956	Swiss-Prot
Mmp1, Oryctolagus cuniculus	Rabbit	P13943	Swiss-Prot
Rana catesbeiana	Bull frog	Q11133	Swiss-Prot
MMP-8 (neutrophil collagenase)			
MMP8 or *CLG1, Homo sapiens*	Human	P22894	Swiss-Prot
CMMP			
Gallus gallus	Chicken		
Enamelysin			
Bos taurus	Bovin	AF009922	GenBank
Sus scrofa	Pig	U54825	GenBank
MMP-7 (matrilysin)			
Mmp7, Mus musculus	Mouse	Q10738	Swiss-Prot
Mmp7, Rattus norvegicus	Rat	P50280	Swiss-Prot
Mmp7, Felis silvestris	Cat	P55032	Swiss-Prot
MMP7, Homo sapiens	Human	P09237	Swiss-Prot
MMP-11 (stromelysin-3)			
MMP11 or *STMY3, Homo sapiens*	Human	P24347	Swiss-Prot
Mmp11, Mus musculus	Mouse	Q02853	Swiss-Prot
Xenopus laevis	Frog	Q11005	Swiss-Prot
Rattus norvegicus	Rat	P97576	TREMBL
MMP-16 (MT3-MMP)			
MMP16 or *MMPX2, Homo sapiens*	Human	P51512	Swiss-Prot
MT3-MMP, Gallus gallus	Chicken	U66463	GenBank

(*continued on next page*)

TABLE 1. (*continued*)

Gene, organism	Common name	Accession #	Database
MMP-14 (MT1-MMP)			
MMP14 or *MMP-X1, Homo sapiens*	Human	P50281	Swiss-Prot
Mmp14 or *MT-MMP, Rattus norvegicus*	Rat	Q10739	Swiss-Prot
Mmp14 or *MT-MMP, Mus musculus*	Mouse	P53690	Swiss-Prot
Mmp14, Oryctolagus cuniculus	Rabbit	Q95220	Swiss-Prot
MMP-15 (MT2-MMP)			
MMP15, Homo sapiens	Human	P51511	Swiss-Prot
MMP-17 (MT4-MMP)			
MMP17, Homo sapiens	Human	X89576	GenBank
MMP-19			
MMP19, Homo sapiens	Human	X92521	GenBank
Plant metalloproteases			
Glycine max	soybean	U63725	GenBank
Arabidopsis thaliana	Mouse ear cress	O04529	TREMBL
Arabidopsis thaliana	Mouse ear cress	E327511	GenBank
Envelysin			
Hemicentrotus pulcherrimus	Sea urchin	AB000719	GenBank
Paracentrotus lividus	Sea urchin	P22757	Swiss-Prot
Zinc protease			
T21D11.1, Caenorhabditis elegans	Nematode	U00038	GenBank
XMMP			
Xenopus laevis	Frog		
Others			
Cynops pyrrhogaster	Newt	D82053	GenBank
Cynops pyrrhogaster	Newt	D82054	GenBank

their similarities and differences (30, 31). We have analyzed the entire sequences, the catalytic domains, and hemopexin-like domains of the 64 members of the MMP family in terms of structures, evolution, and interactions with substrates and inhibitors. We also provide an analysis of the structural zinc binding site and the calcium binding site. These are the first comprehensive analyses of this important family of enzymes, and provide fundamental information on evolution and properties of MMPs.

EXPERIMENTAL PROCEDURES

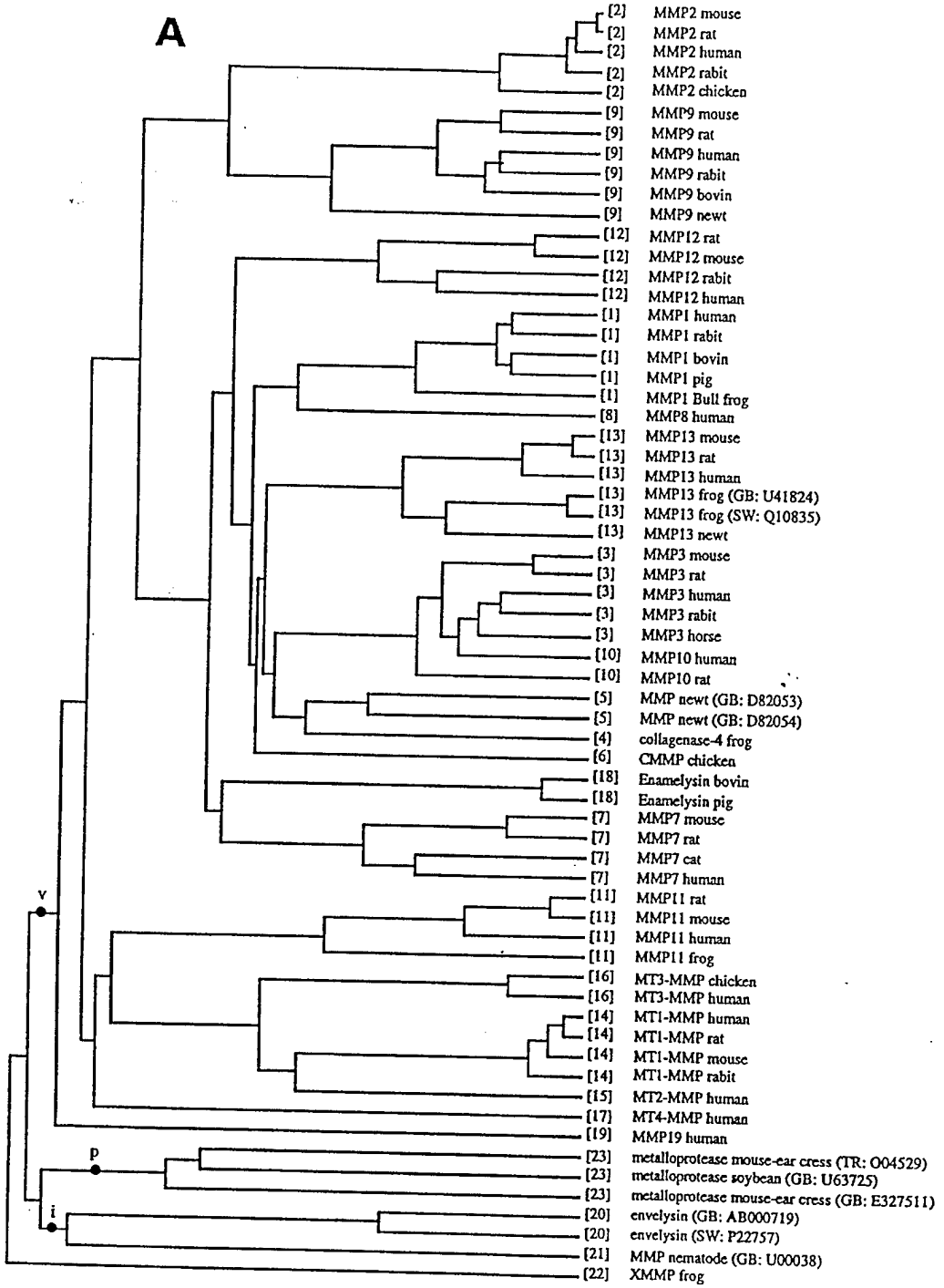
Amino acid sequences of MMPs were obtained from the GenBank, TREMBL, and Swiss-Prot data banks. The sources for all sequences are given in Table 1. A total of 64 different MMPs were used for this analysis [human enamelysin and the mmp20 gene product with GenBank accession # AJ003147 were not included because they were reported after the completion of our analysis. However, we conducted a separate multiple-sequence analysis with the 64 sequences and the two new MMP sequences after preparing this review. This analysis showed that the clustering pattern was identical to that reported in this article. Human enamelysin grouped with other enamelysins, and the mmp20 gene product showed the closest homology to MT4-MMP]. The multiple-sequence alignments were performed using the program PileUp from the Wisconsin package version 9. Four human MMPs [fibroblast (MMP-1, 1cgl) (26) and neutrophil (MMP-8, 1mnc) (27) collagenases, matrilysin (MMP-7, 1mmq) (28), and stromelysin-1 (MMP-3, 1slm) (29)] have recently been crystallized and their coordinates are available. We used structural information to predict the 3-dimensional structures for the homologous metalloproteinases using the program COMPOSER (Tripos Associates, Inc., St. Louis, Mo.).

We recently reported the computational 3-dimensional models for the catalytic domains of MMP-2 and MMP-9 (30) and also predicted the folding of an additional 17 representative MMP enzymes for which such information has been lacking (31). These 17 proteins are human MMP-10, MMP-11, MMP-12, MMP-13, MT1-MMP (MMP-14), MT2-MMP (MMP-15), MT3-MMP (MMP-16), MT4-MMP (MMP-17), MMP-19 (also referred to as MMP-18), pig enamelysin, sea urchin envelysin, stromelysin-like MMP from newt, collagenase-4 from frog, nematode MMP, chicken CMMP, frog XMMP, and MMP from mouse ear cress.

RESULTS AND DISCUSSION

Multiple-sequence analysis for the 64 MMPs was conducted on three different sets of data. In the first set,

the complete sequences, including the signal and propeptide regions, were used (**Fig. 2A**). A simplified schematic presentation of Fig. 2*A* is shown as Fig. 2*B*. The entire sequences were used to understand the overall evolutionary pathways for diversification. Evolution occurs via separate events of sequence modification of the entire gene such as point mutations, insertions, deletions, gene splitting, and fusions. Mu-



(*continued on next page*)

Figure 2. *A)* Dendrogram for the multiple-sequence analysis of the complete amino acid sequences for 64 MMPs. Numbers in brackets represent the classes. The letter 'v' indicates the branch for vertebrate MMPs, 'i' indicates invertebrate MMPs, and 'p' indicates the plant MMPs. *B)* Simplified schematic of the dendrogram shown in panel *A*. The numbers within circles represent chromosomal origin for human MMPs, when available.

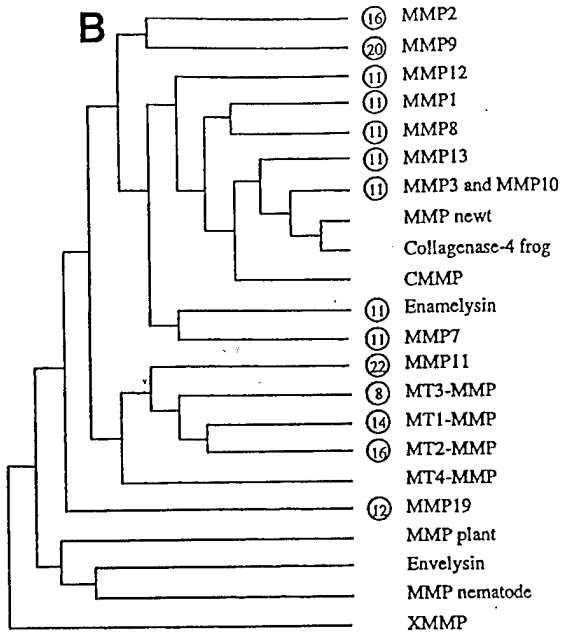


Figure 2. (*continued*)

tations occur at random regardless of the function of the region of the gene or the domain in order to increase diversity throughout the genes in question, which is why this sequence analysis on the entire protein is informative. The second and third sets for multiple-sequence analyses were carried out on sequence stretches corresponding to the catalytic and hemopexin-like domains of the MMPs (**Fig. 3** and **Fig. 4**, respectively). These two sets were chosen to study the different evolutionary pressures on these specific domains en route to diversification. Mutations that impair catalytic ability would not be selected. In contrast, substrate specificities for the majority of the MMPs are believed to have been determined by the interactions of protein substrates and TIMPs with either the hemopexin-like domain or the gelatin binding domain (in the case of gelatinases) (32, 33). Hence, evolution of these domains may reflect evolution of substrate specificities and interactions with TIMPs. Therefore, comparison of the data for analyses of the catalytic domain and hemopexin-like domain with those of the complete sequences would provide information on whether the assembly of the genes encoding these domains occurred at a later time than the differentiation of MMPs into separate subfamilies.

Analysis of the entire sequences and the separate analyses of the catalytic and hemopexin-like domains are quite revealing. Analysis of the entire sequences and of the catalytic domains gave rise to 22 distinct subfamilies of MMPs. The number for each cluster is given in brackets and is arbitrary. We tried to correlate the number for the clusters with those for the given MMP when possible. For example, cluster 1 is given to MMP-1. Each analysis produced clusters of enzymes that were individually comprised of MMPs of a given type. The only exceptions were MMP-3 and MMP-10 (stromelysins-1 and -2, respectively), which clustered into one subfamily (Figs. 2 and 3). However, these stromelysins formed two independent groups in the alignment of the hemopexin-like domains, giving rise to a total of 23 distinct subfamilies of MMPs for this analysis. **Figure 5** provides the alignment for the amino acid sequences of the 23 representative enzymes using the PileUp program. This alignment was made consistent with the alignments found by the program COMPOSER, which takes into account the predicted folded structures of the proteins. COMPOSER assigns different contributions for the 'structurally conserved regions', such as elements of well-defined secondary structure, for example, β-strands and α-helices, as opposed to loops, which are considered to be variable areas. In contrast, PileUp does not differentiate between secondary structure elements as does COMPOSER, but rather performs the alignment based solely on sequence homology of amino acids. Therefore, sequence alignments found by COMPOSER and PileUp are based on different principles, yet they have the potential to complement each other. We have edited the result of the PileUp alignment from the insight gained by the COMPOSER analysis, since the 3-dimensional structural information would enhance reliability of this type of analysis considerably. Major protein structural blocks, β-strands and α-helices, have a greater tendency to be conserved during the evolutionary process than do mere sequences of amino acids.

Multiple-sequence analysis of the complete sequences of MMPs

The 22 major subfamilies of MMPs, as discerned from the multiple-sequence analysis and the dendrogram for the 64 MMPs, are shown in Fig. 2*A*, with a simplified version depicted in Fig. 2*B*. The chromosome locations for human MMPs are shown (in circles) in Fig. 2*B* when available. Members of each subfamily generally display similar substrate profiles in different organisms. MMP-3 and MMP-10 clustered in one group, and show almost identical properties and substrate profiles (8). Genes of the more closely clustered human MMPs all seem to have originated from the same chromosome 11, reflecting the evolutionary process suggested in Fig. 2*B*.

It has been proposed that the origin of MMPs could be traced to before the emergence of vertebrates from invertebrates (34, 35). Recent sequencing of three plant MMPs found these enzymes to be homologous to MMPs from vertebrate and invertebrate origins, indicating that this evolutionary process is even more ancient than previously appreciated

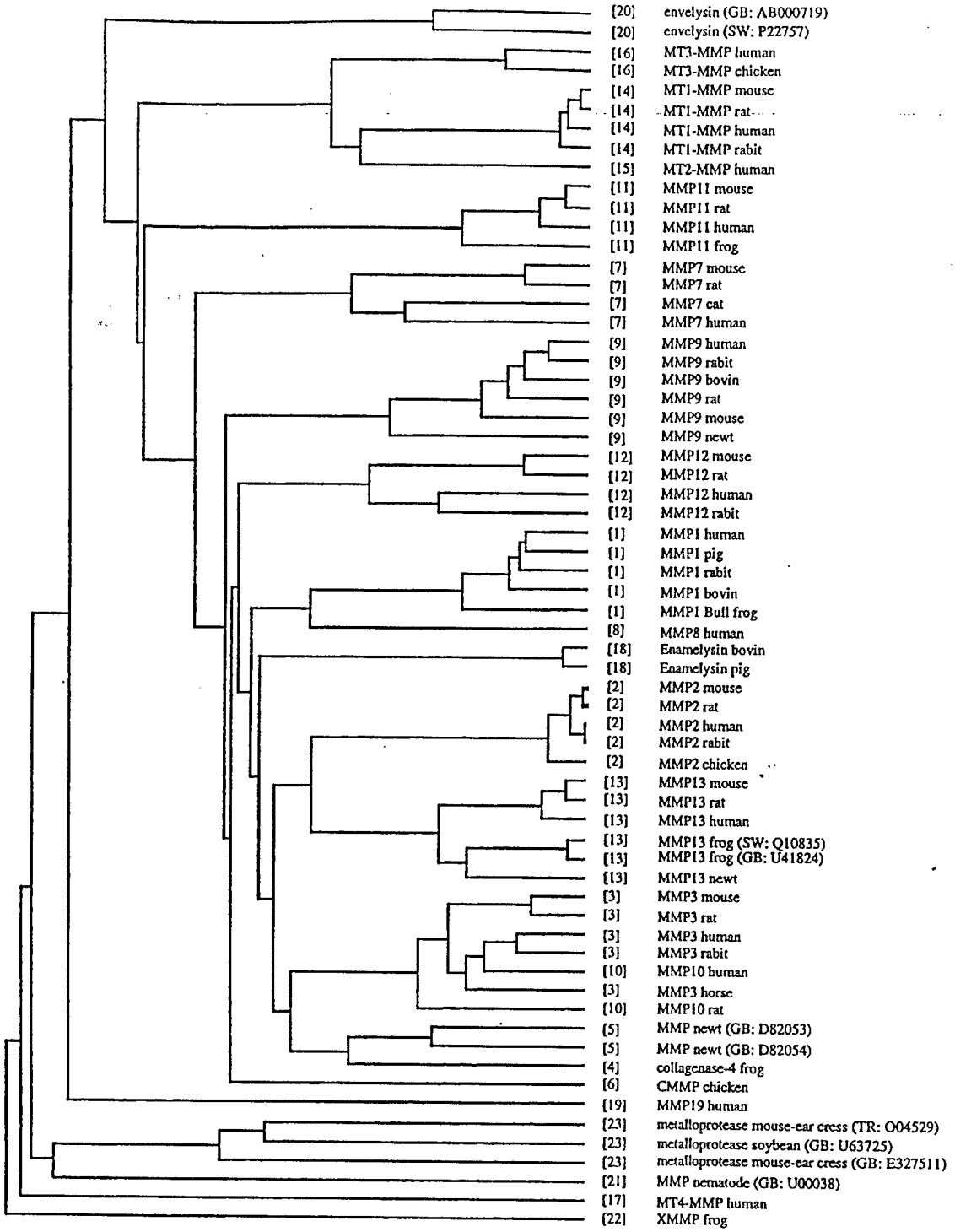


Figure 3. Dendrogram for the multiple-sequence analysis for the sequence of amino acids in the catalytic domains for 64 MMPs. The numbers in brackets represent the classes.

(Table 1) (36). Furthermore, the discovery of a peptide sequence for a metal-containing enzyme from *B. fragilis* suggests that MMPs may be more ancient yet (37), given that bacteria have been around for longer than 3.5 billion years. The three main branches of the dendrogram in Fig. 2*A* give rise to the lines that lead to enzymes from the vertebrate (depicted as the 'v' branch), invertebrate (depicted as the 'i' branch), and plant (depicted as the 'p' branch) MMPs, respectively. The only exception is the XMMP from

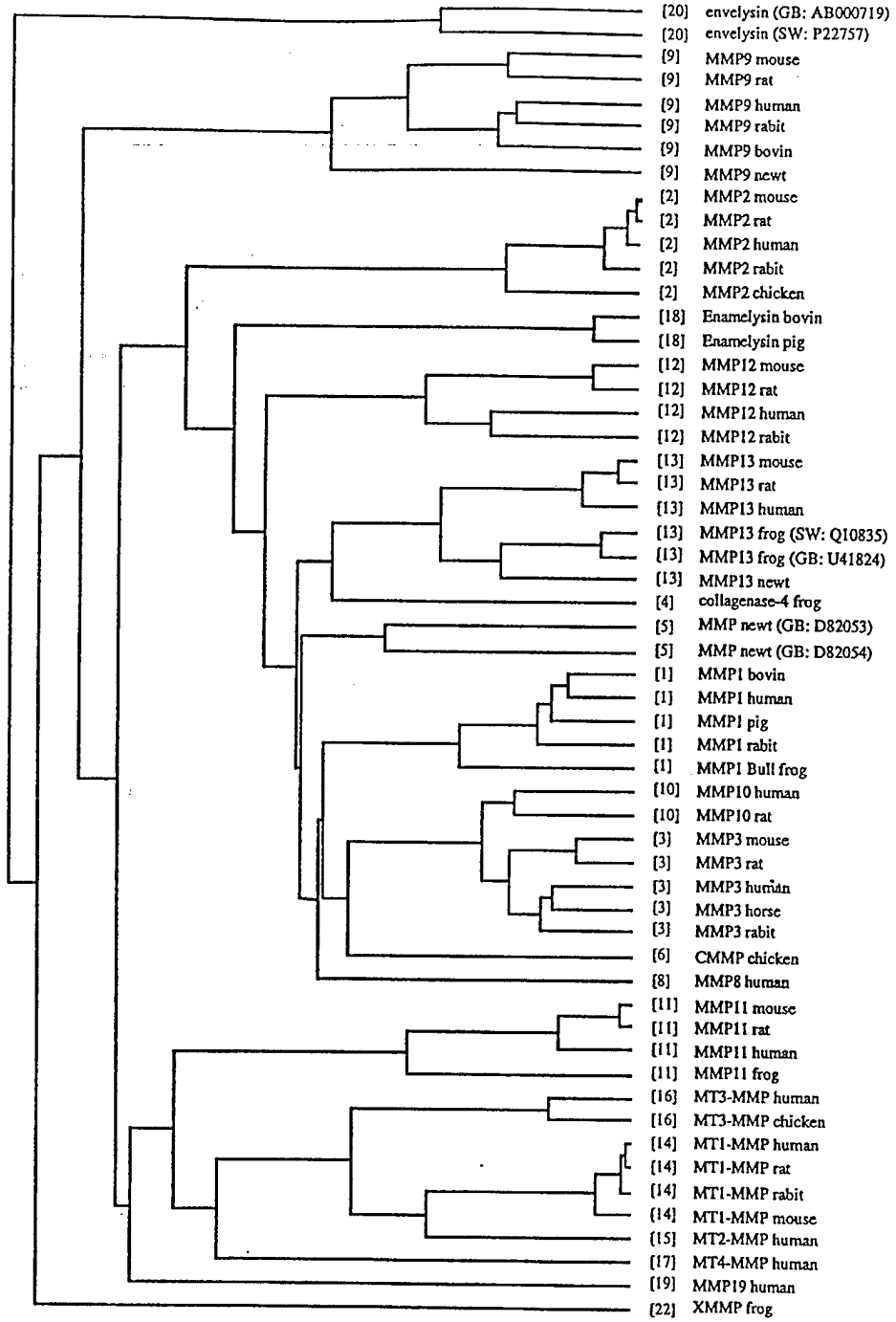


Figure 4. Dendrogram for the multiple-sequence analysis of the hemopexin-like domain sequences for 64 MMPs. Numbers in parentheses represent the class.

Xenopus (38). Plant and invertebrate MMPs show a stronger relationship among themselves than to the vertebrate MMPs (Fig. 2*A*). Thus, the vertebrate enzymes are more remotely related to plant and invertebrate MMPs. Although one would expect that the plant MMPs are the most ancient, and therefore the enzymes least related to all other MMPs, our analysis shows that the frog XMMP is the enzyme least related to all other MMPs, itself forming a separate group. This indicates that XMMP either represents a separate yet unidentified group of MMPs or is the last extant member of a primordial MMP. The first possibility seems to be more likely in our view, because XMMP possesses a hemopexin-like domain that is absent in plant MMPs and the nematode enzyme. Indeed, the plant and nematode MMPs have the sim-

plest domain structures. It may be suggested that at some point during evolution, the genes encoding the primordial MMP and hemopexin-like domain joined together, and XMMP probably originated after this event. Based on our analysis of Fig. 2*A*, we see that the sequence of MMP-7 fits well into the alignment of the entire sequences. This is also true for the sequence of the catalytic domains (vide infra), indicating that this enzyme has not existed as an evolutionary exception, or oddity, by not having the hemopexin-like domain. In light of what appears to be an independent, and perhaps parallel, evolution for MMP-7, the lack of hemopexin-like domain in matrilysin represents a deletion of this domain during evolution (Fig. 1*A*).

One can see four major subgroups within the vertebrate branch (Fig. 2*A*). One is formed by MMP-19 itself (some refer to this as MMP-18), which is least related to other vertebrate MMPs, besides XMMP. The next branch is that of MMP-11 (stromelysin-3) clustered together with the MT-MMPs. MT4-MMP (MMP-17) shares less similarity to other members of this branch of membrane-type enzymes, and the rest of this subgroup is more closely related to MMP-11 than to MT4-MMP. Both MMP-11 and the MT-MMPs possess the furin recognition site and can be activated by furin-like convertases (11, 39). The third branch is comprised exclusively of gelatinases, whereas the last branch is made up of all remaining vertebrate MMPs.

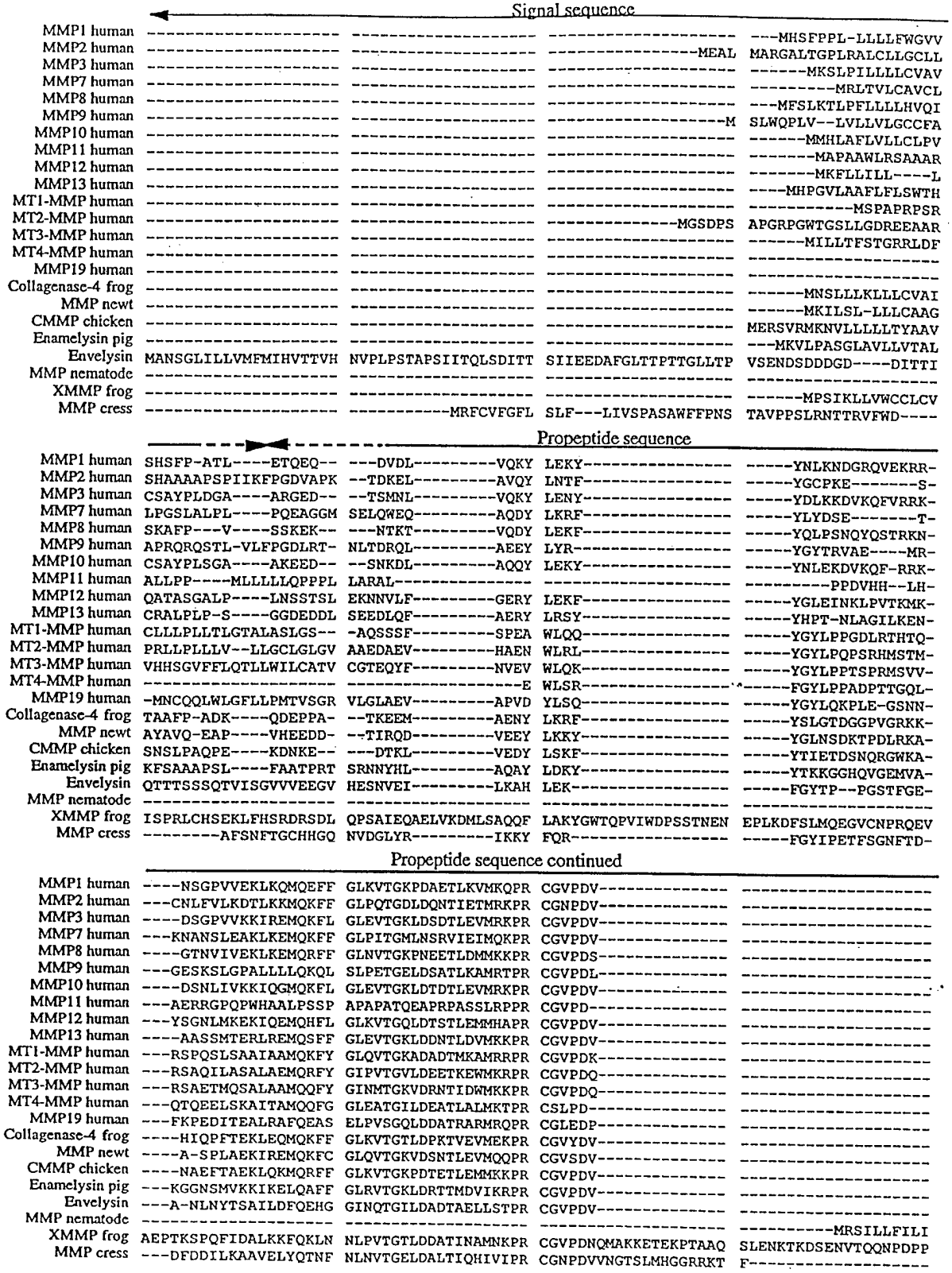
Figures 2–4 reveal 22–23 distinct MMPs subfamilies, sequence similarities of which were preserved among various organisms during evolution. It is likely that one would find counterparts to each MMP in various vertebrate organisms. The entire amino acid sequence alignment of the 23 representative members of these subfamilies is shown in Fig. 5. Human enzymes have been selected in all cases when available. The simplified domain structures of these 23 MMPs are shown in the schematic of Fig. 1*B*. Consistent with earlier knowledge, the analysis depicted in Fig. 5 and the alignment produced by the PileUp program revealed that all MMPs, except that from nematode, have signal sequences, followed by the propeptide region containing the characteristic motif called the 'cysteine switch' (40). The common pattern for this cysteine switch is [PSL]-[RT]-C-[GS]-[VNL]-[PASYE]-**D**, where boldface letters designate the most commonly found amino acids. In addition, MMP-11 and the MT-MMPs (except MT4-MMP) contain a RXKR motif in the propeptide that has been postulated to represent a cleavage site for furin-like enzymes (11, 13, 39). This cleavage pattern also appears in nematode MMP (two repeats) and in XMMP, as seen in the sequence alignment. MT4-MMP and the mmp20 gene product have a variation of the furin-like recognition sequence as RRRR (instead of RXKR) at this position.

The sequence of the nematode MMP seems to be considerably shorter at the amino terminus (24), which either reflects that the amino-terminal portion of the enzyme was not sequenced entirely or that this enzyme has a totally different mechanism for activation. The amino-terminal portion of the nematode MMP that precedes the catalytic domain has only 48 amino acids, compared to 100–200 amino acids in other MMPs. The propeptide sequence forms a cap over the active site of this MMP, with the critical cysteine residue providing the fourth coordination to the catalytic zinc ion, as seen in the crystal structure of stromelysin-1 (41). We also see a conserved signature for the catalytic zinc ion binding site of MMPs, with the consensus pattern containing a so-called methionine turn (40, 42), [VAIT]-[AG]-[ATV]-*H*-E-[FLIV]-G-*H*-[ALMSV]-[LIM]-G-[LM]-X-*H*-[SITV]-X(5)-[LAFIV]-*M*, where X denotes any residue and X(5) means there are five residues between the flanking sites. The three histidines shown in italics chelate the catalytic zinc ion, and the methionine (also depicted in italics) is located underneath the cavity formed by these histidines, providing increased hydrophobicity in this area to enhance zinc binding ability of histidines (42).

Multiple-sequence analysis of regions encoding the catalytic domain of MMPs

Figure 3 shows the multiple-sequence analysis carried out for the regions encoding exclusively the catalytic domains. Again one sees 22 distinct clusters, with MMP-3 and MMP-10 in one group. We no longer see three clearly delineated (separate) large groups for invertebrate, vertebrate, and plant MMPs, although there appears to be some tendency for MMPs from these sources to group together. The catalytic domains of the MT-MMPs do not usually appear to be closely related to one another (for example, position of MT4-MMP in Fig. 3). An unexpected finding was

Figure 5. Amino acid sequence alignment for the 23 representative MMPs. The pound sign '#' marks the residues important for binding of the zinc ion in the active site of MMPs. The symbol '@' denotes the residues coordinated to the structural zinc ion, '%' marks residues that coordinate to the calcium ion by their side chain functions, and the letter 'B' labels residues that contribute their main-chain carbonyl moieties for coordination to the calcium ion. The letters 'X', 'Y', and 'Z' denote the structurally variable loops important for substrate binding to the catalytic domain. The residues marked by 'J' provide elements of their backbone to anchor the substrate. *continued on next page*



(*continued on next page*)

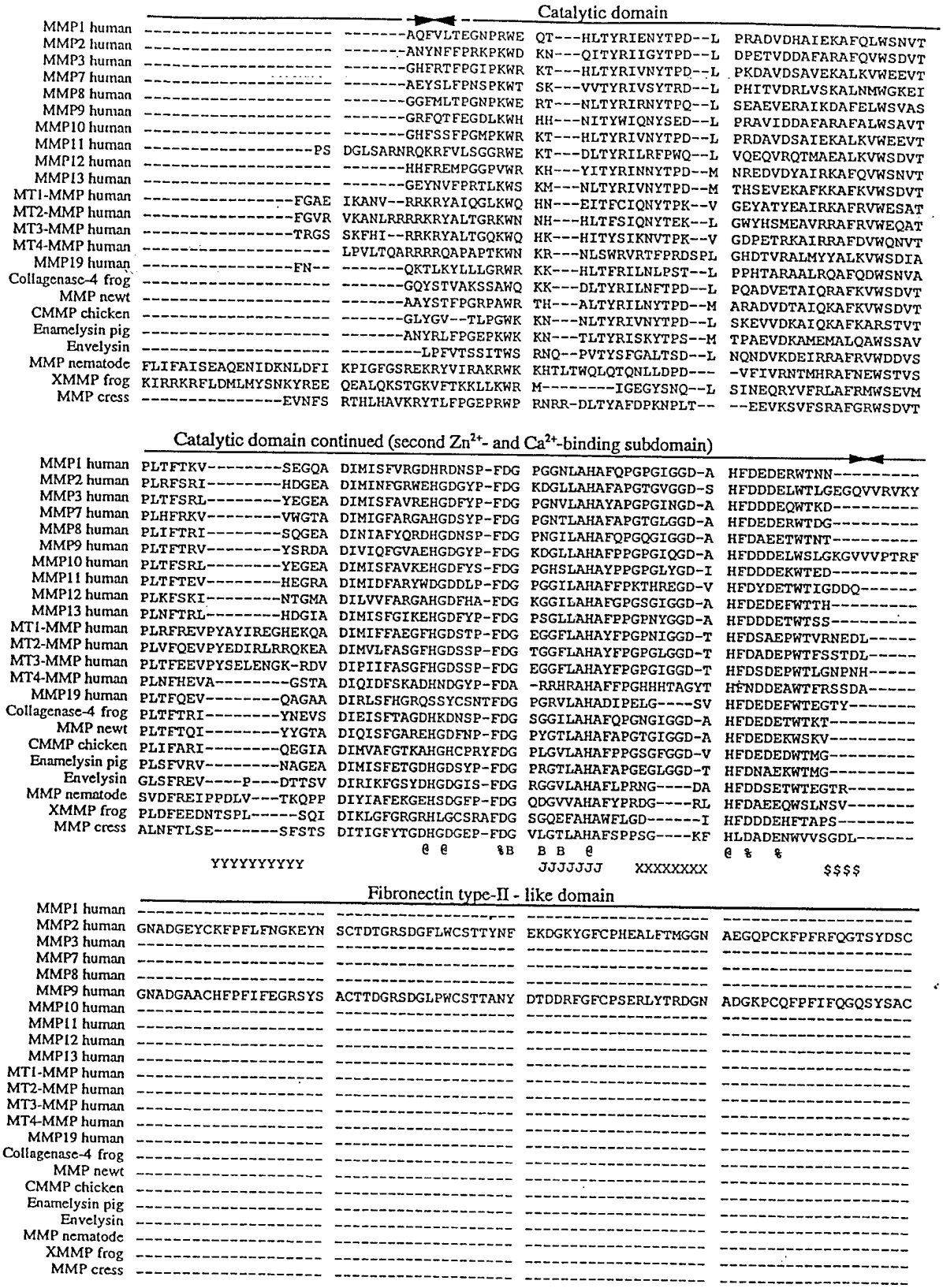


Figure 5. (*continued*)

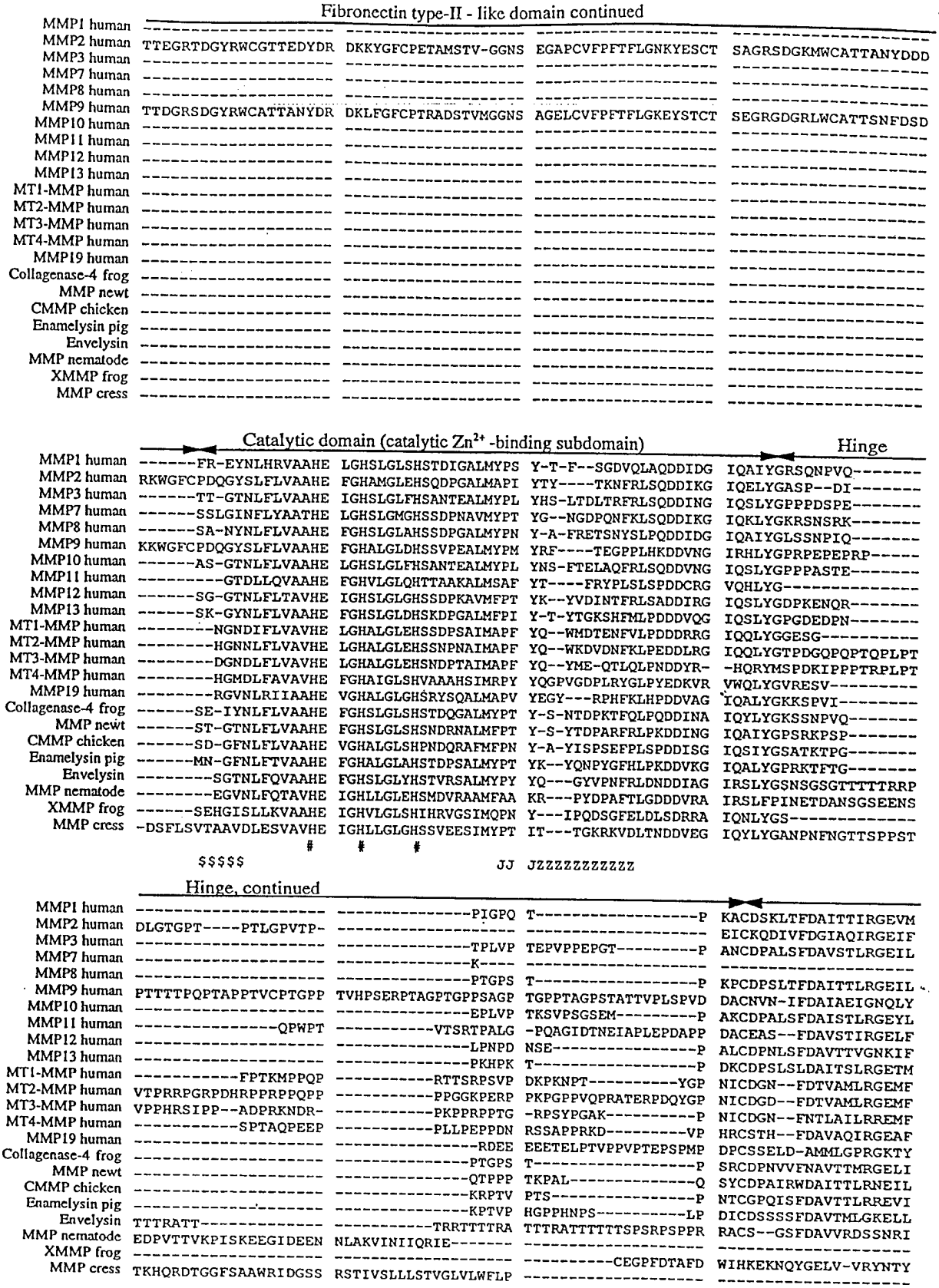


Figure 5. (*continued*)

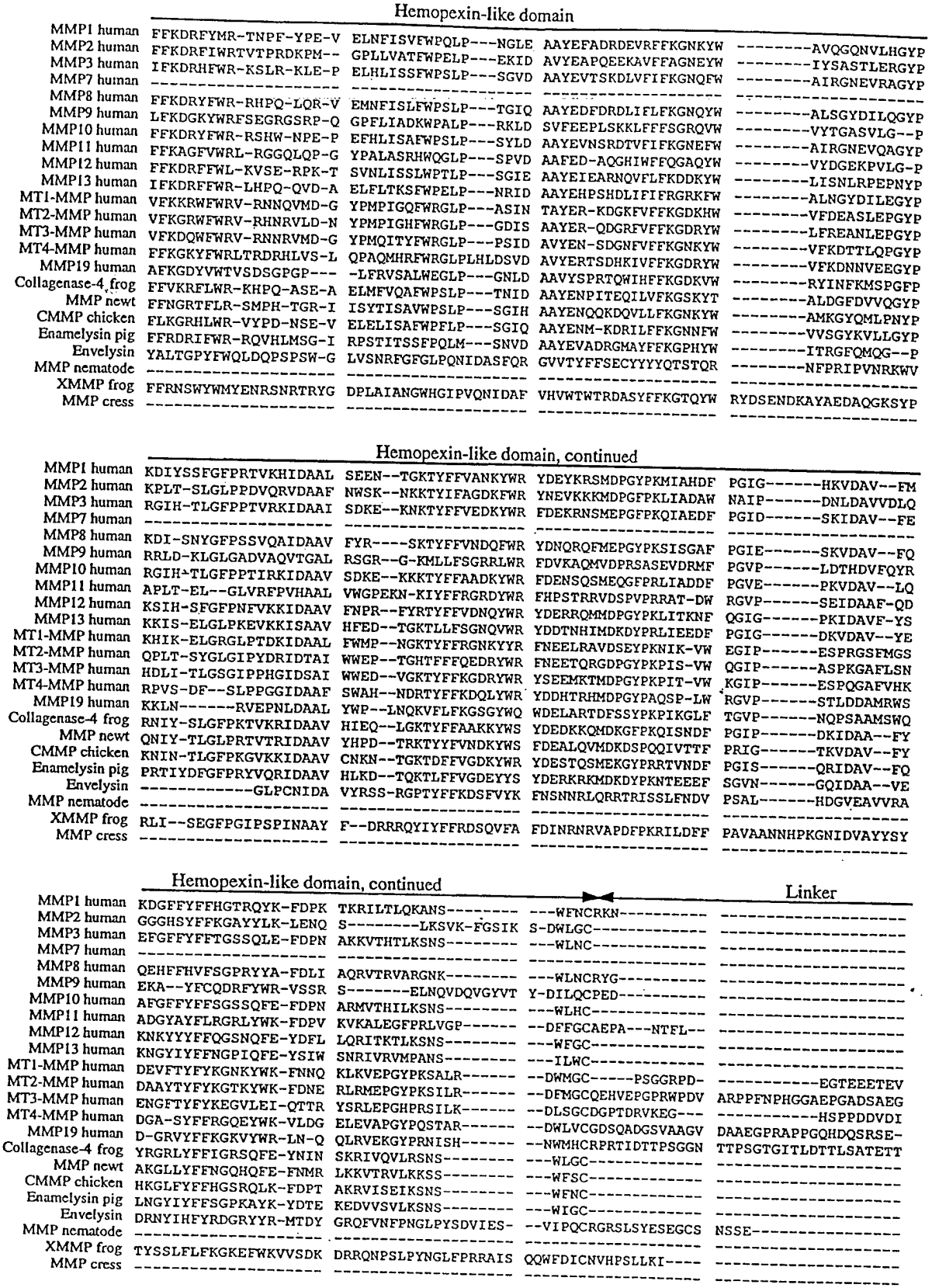


Figure 5. (*continued*)

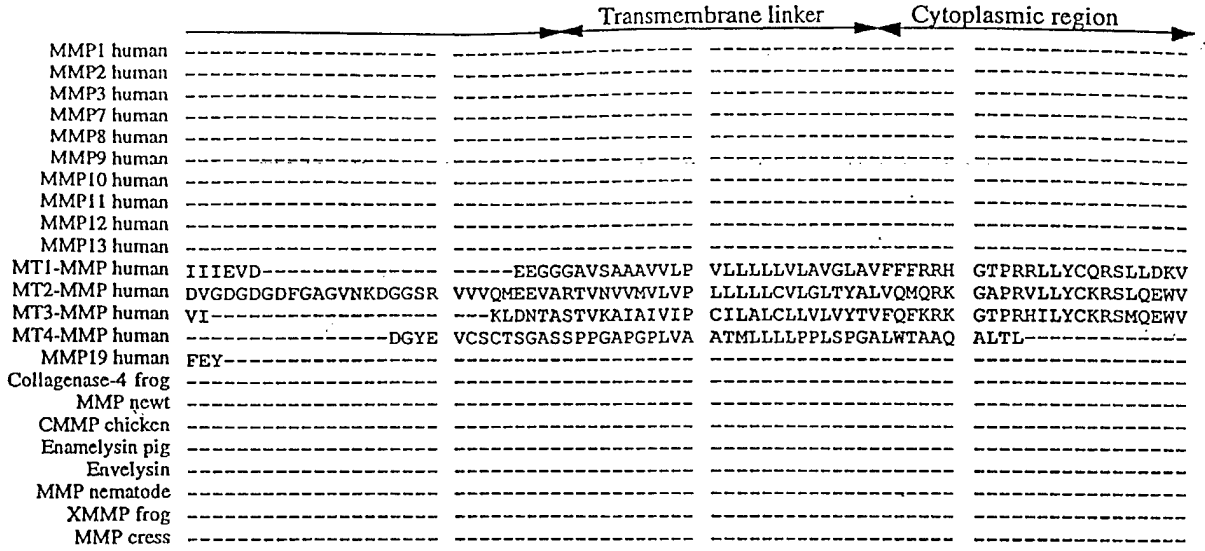


Figure 5. (*continued*)

that the catalytic domains of MMP-2 and MMP-9 (gelatinases) are not clustered together. Also, unlike sequence analysis for the entire enzymes, sequences for the catalytic domains did not sequester according to the location of the respective genes within the given chromosome. These observations collectively argue for the fact that the catalytic domains of these MMPs likely evolved in parallel, indicating that the selection pressure for the catalytic domain was distinct in the course of the diversification of this family of enzymes.

Three-dimensional structures of the catalytic domains of the 23 representative MMPs

The unique features of 3-dimensional structures of these enzymes were studied by developing computational models for the 23 catalytic domains. The crystal structures of the catalytic domains of four MMPs (MMP-1, MMP-3, MMP-7, and MMP-8) have been elucidated (26–29). We modeled the catalytic domain structures of the remaining 19 representative enzymes on the basis of the similar fold of the proteins that have been crystallized (30, 31; coordinates for the structures of the modeled MMP catalytic domains can be obtained from our group web page: http://sun2.science.wayne.edu/~somgroup). Closer examination of the structures of the catalytic domains revealed that a conserved aspartic acid is found in the vicinity of the methionine turn, the side chain of which is buried inside the core of the domain. The only variation to this pattern is seen in the rabbit MT1-MMP, which has a glutamic acid in this position. The two side chain oxygens of the aspartic acid form two critical structural hydrogen bonds to the backbone amides, one with the methionine of the methionine turn and another with the residue preceding the methionine. The conservation of this pattern/motif in all 64 known MMPs argues that these three histidines are absolutely required structural elements for the precise positioning of the catalytic zinc ion for effective catalysis. These three amino acids are marked with the pound sign (#) in Fig. 5.

We investigated the nature of two additional zinc and calcium binding sites formed by β-strands and turns in the proximity of the catalytic zinc in all the MMPs (31). The 64 MMPs developed at least four different ways to bind to this structural zinc ion. For the majority of MMPs (60 enzymes representing 20 subfamilies of the total of 23 identified by us), this site is provided by the side chains of an aspartic acid and three histidines (marked as an in Fig. 5). The signature for the binding site of the structural zinc and the calcium ion in the 60 MMPs is H_δ-[GN]-***D***-X(2)-[PAS]-F-***D***-[GA]-X(4)-[LIRV]-[AG]-H_δ-[AV]-[FYS]-P-X(5,7,9)-H_ε-[FL]-***D***-X(2)-**E**-X-**W**. The letters in bold italics represent residues that provide side chains for coordination to the structural zinc and calcium; X(5, 7, 9) indicates five, seven, or nine variable residues in between the flanking sites. The . . .FYS. . . region in this pattern is close to the coordinated histidines, a result of the enzyme fold. The presence of these hydrophobic residues creates an increased hydrophobic environment and enhances the binding affinity for the metal ions. The human MMP-11 has an aspartic acid instead of one of His_δ, which is different from the MMP-11 from rabbit and mouse, which still have a histidine at this position, suggesting that MMP-11 (comparing those from human, rabbit, and mouse) does not possess a unique motif for binding of the structural zinc. The chicken CMMP, the frog XMMP, and the human MMP-19 form a separate subfamily in the multiple-sequence

dendrograms and each has a unique motif for coordination to the structural zinc ion (31). The chicken CMMP has a histidine present where the majority of MMPs have an aspartic acid coordinated to the second zinc ion. A cysteine residue (Cys^{174}) adjacent to the fourth coordinated histidine implicates two possibilities for the binding mode of the structural zinc in CMMP. This cysteine may either provide a site for protein dimerization or another binding mode for the second zinc ion in CMMP (three original histidines and the cysteine coordinated to the zinc ion). This cysteine residue is also present in the sequence of the frog XMMP at the same position. The structural zinc coordination in MMP-19 is seen with two histidines at usual positions and one cysteine at the corresponding position as discussed for CMMP and XMMP, and the third histidine four residues toward the amino terminus. It has a noncoordination serine at the position where most MMPs have the third histidine (Fig. 5). The position of the third histidine is on a different β-strand, and the orientation of the side chain for the fourth coordination site is perfectly acceptable. Hence, the variations noted in the coordination to structural zinc ion may be indicative of the different outcomes for selection of novel enzymic activities. Furthermore, the dendrogram shown in Fig. 3 indicates that these variations in the zinc binding motif came about as a consequence of independent, unrelated evolutionary processes.

Calcium ion coordinates with six elements in the 3-dimensional structures of catalytic domains of MMPs, which come into close proximity for coordination as an octahedron. In all 64 MMPs, three of these elements, which are the side chains of amino acid residues, are conserved. These are two aspartic acids and one glutamic acid (marked by '%' in Fig. 5). A minor variation is seen only for the human MT4-MMP, which has an asparagine residue instead of the second aspartic acid. The other three calcium ligands are provided by the backbone carbonyl oxygens of three residues within a turn. Positions are marked by the letter 'B' in Fig. 5. The spatial location of the catalytic zinc ion, structural zinc ion, and the calcium ion are shown in Fig. 6.

Multiple-sequence analysis of regions encoding the hemopexin-like domain of MMPs

Figure 4 shows the results of multiple-sequence alignment for the hemopexin-like domains by themselves. This domain is absent in MMP-7 (matrilysin), in all known plant MMPs, and in the nematode MMP. The sequences of the hemopexin-like domains of invertebrate MMPs are the least related to all other MMPs, which indicates their ancient origin. In some MMPs, the hemopexin-like domains have been shown to facilitate binding and denaturation of the macromolecular substrates; it would be interesting to correlate clustering schemes for the hemopexin-like domains of MMPs to their substrate specificities. Because of the multiplicity of the known substrates for some of these enzymes and, in contrast, the paucity of any information on substrates for other MMPs, it is extremely difficult to draw any substantial conclusions on this issue. On the other hand, one cannot help but notice the diversification of the hemopexin-like domains seen in these proteins; insofar as this domain is clearly linked to the issue of substrate specificity, the diversity in this domain for the various MMPs indicates different evolutionary tangents pursued by these functionally distinct enzymes.

Murphy and Knäuper (32) recently reviewed the role of the hemopexin-like domains in relation to the substrate specificities and activities of various MMPs. It was suggested that the hemopexin-like domains mediate binding of MMP-1, MMP-8, MMP-13, and MMP-3 to collagen and that the complex participates in the cleavage of triple helical collagen. In the case of MMP-2 and MMP-9, the hemopexin-like domain is important for interactions with TIMPs, although the high degree of sequence similarity and the likely structures of the hemopexin-like domains of the gelatinases suggest there is a high degree of specificity in the binding of TIMPs to the latent forms of these enzymes. For example, TIMP-1 binds exclusively to latent MMP-9 (K_d ~35 nM), whereas TIMP-2 binds to latent MMP-2 (K_d ~5 nM) (43). We have recently shown a biphasic binding of TIMP-1 and TIMP-2 to the latent forms of MMP-9 and MMP-2, respectively, with the hemopexin-like domain representing the high-affinity binding site (43). Removal of the hemopexin-like domain of MMP-2 decreases the affinity of TIMP-2 for the active site without significantly affecting enzymatic activity. TIMP-1, which efficiently inhibits the active form of MMP-2, does not bind to a carboxy-terminally truncated MMP-2 form, demonstrating the importance of the hemopexin-like domain in interactions of TIMP-1 with active MMP-2 (43). The hemopexin-like domain of MMP-2 has also been shown to play a role in zymogen activation by MT1-MMP (44, 45). It has also been suggested that the hemopexin-like domain of MMP-2 plays a role in the binding of the enzyme to integrin $\alpha_v\beta_3$ (46, 47), a process that may facilitate localization of MMP-2 on the cell surface. The alignment of the hemopexin-like domains shows that MMP-2 and MMP-9 fall into two different clusters. The hemopexin-like domain of MMP-9 diverges at a higher hierarchial level, implicating that it is somewhat different from that of MMP-2 even though both are involved in TIMP binding. In contrast to gelatinases, there is not enough biochemical data on the various roles of hemopexin-like domains for all known MMPs, which may be different in each case. The interesting diversification of this domain and its effect on the functions of various MMPs can also be seen in the dendrogram of Fig. 4.

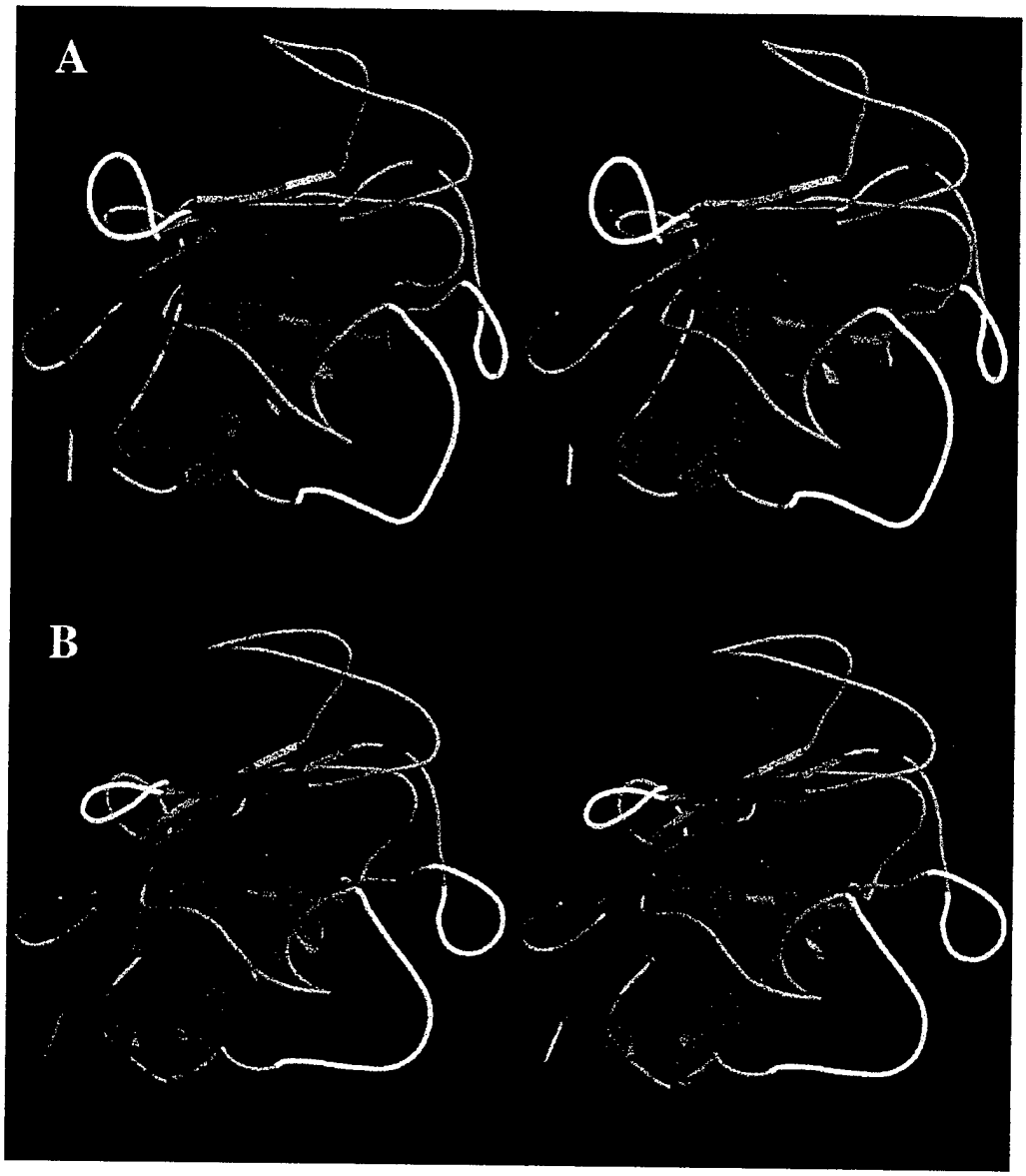


Figure 6. Ribbon drawing of the modeled 3-dimensional structures for the catalytic domains of MMP-12 (*A*) and MMP-19 (*B*). Red spheres represent the zinc ions and the green sphere represents the calcium ion. Catalytic zinc ion is located at the center of the catalytic domain and structural zinc is at the 12 o'clock position. The variable loops at 10, 5, and 3 o'clock (designated by 'X', 'Z', and '$' in Fig. 5, respectively) are shown in white. The two models are similar structurally, except for the variable regions (shown in white). Figures were prepared using MOLSCRIPT and Raster3D rendering programs (57, 58).

With the exception of envelysin and XMMP, all other hemopexin-like domains fall under one cluster. The hemopexin-like domain of MMP-9 formed a cluster by itself (vide infra), whereas those of MMP-11, MMP-19, and the MT-MMPs formed a separate cluster; all remaining MMPs constitute an additional cluster.

The hemopexin-like domain of envelysin forms a separate cluster from those of the other MMPs. Envelysin degrades the protective fertilization envelope, a complex of glycoproteins, releasing the embryos of sea urchin, although the individual glycoproteins have not yet been identified. Envelysin is also known to hydrolyze small peptides like substance P, oxidized insulin B, and collagenase substrate-like small peptide (48). In a recent study, it was also shown that the hemopexin-like domain of envelysin determines substrate specificity for this enzyme (24). The substrate specificity of envelysin is believed to be similar to that of stromelysin-1, which also degrades the fertilization envelope proteins of greater than 100 kDa. From the position of envelysin in the dendrogram as well as from the available functional data on this enzyme, it would appear that the hemopexin-like domain of this MMP is distinct from the rest and diverged early from those of the other MMPs.

Enamelysin is detected during the development of the enamel matrix and is expressed specifically in the enamel tissue. According to this report as well as one by other investigators (49), enamelysin forms a separate subfamily of MMPs. It is difficult to correlate its substrate specificity to the hemopexin-like domain alignment from the limited biochemical data available. However, its position suggests that it may be quite different from other collagenases. Recently, MMP-20 has been sequenced and its gene has been mapped to chromosome 11 (21). Our analysis of the complete sequences of MMPs (without human enamelysin) showed that pig enamelysin and other MMPs that are mapped together to chromosome 11 cluster together (Fig. 2*B*). In GenBank, another human metalloproteinase that is a product of the gene mmp20 has recently been reported (accession number AJ003144); this enzyme is mapped to chromosome 16. Only 183 amino acids are reported in its sequence; it probably is not sequenced completely and is not included in our current analysis (22).

The hemopexin-like domain of MMP-12 shows clustering by itself into a separate subfamily in close relation to enamelysin. Thus, the hemopexin-like domain of MMP-12 appears to have diverged into a separate subfamily (i.e., specialized) earlier than the hemopexin-like domains of the stromelysins and collagenases (MMP-1, MMP-3, MMP-8, MMP-10, and MMP-13, the exception being MMP-11), underscoring the role of MMP-12 as an 'elastase'. Despite the lack of a hemopexin-like domain, MMP-7 possesses a substrate preference similar to that of MMP-12, suggesting that the influence of the hemopexin-like domain on substrate interactions is limited, and other sites may play a role in determining substrate specificity. Indeed, comparison of the entire sequences of MMP-7 and MMP-12 (Fig. 2) shows that the clusters of these enzymes are equally remote from other MMPs, which, to put it another way, are equally distinct from other collagenases, gelatinases, stromelysins, and MT-MMPs.

MMP-13 possesses a substrate specificity that is broader than that of the other collagenases such as MMP-1 and MMP-8. The hemopexin-like domain of MMP-13 distinguishes itself by clustering into a separate subfamily, diverged at a higher level of hierarchy than MMP-1 and MMP-8. The frog collagenase-4 cleaves collagen type I, similar to MMP-1, and also possesses a weak gelatinolytic activity. This enzyme has been classified as a different type of collagenase due to its characteristic cleavage pattern of gelatin (50). This is supported by analysis of the sequences of hemopexin-like domains presented here, where the hemopexin-like domain of collagenase-4 clusters differently than those of other collagenases.

The newt MMPs have not been studied extensively to define their substrate specificities. However, in a sequence comparison study, these were classified as 'stromelysin type' (51). In our analysis, the newt MMPs also fell into a subfamily comprised of collagenases and stromelysins (MMP-13, MMP-1, MMP-10, CMMP, and MMP-8). MMP-1, MMP-10, MMP-3, MMP-8, and CMMP form closely akin, independent clusters, indicating that they are related, yet possess differentiated substrate specificities.

The hemopexin-like domain of MMP-11 (stromelysin-3), along with that of MMP-19 and MT-MMPs, diverged and formed a separate subfamily. MMP-11 is the first MMP reported to be activated intracellularly by means of a furin-like convertase and has been shown to be unable to cleave any of the major extracellular matrix components like other collagenases and stromelysins (52). The hemopexin-like domain of MMP-11 clustered separately in our analysis, suggesting distinct functional properties. Analysis of the role of the hemopexin-like domains and their relation to substrate specificities, if any, for the MT-MMP subfamily is more complex due to the presence of the transmembrane domain (vide infra). The hemopexin-like domains of all four known MT-MMPs form a subfamily of their own. Though there is not enough biochemical data on MT-MMPs, from the alignment of the hemopexin-like domains one can say that these domains have diverged to a significant extent from those of the other MMPs (except MMP-11), suggesting that their functional roles are different from the rest. XMMP has not been studied extensively for natural substrates. However, the results of the alignment of the complete sequence and the hemopexin-like domain of XMMP suggest that the substrate profile of XMMP may also be different from those of the other MMPs.

General folding of the catalytic domains of MMPs

In a previous study of the structural aspects of MMP-2 and MMP-9 (30), we investigated the binding modes of peptide substrates in the active sites of six MMPs (four crystallized and two modeled). Residues marked by the letter 'J' in Fig. 5 provide the anchoring interaction to the backbone elements of a potential substrate. The general structural comparison of the 23 representative MMPs revealed four areas of topological variability in the catalytic domains of the 64 MMPs. These areas are formed by four loops, three of which are located in the vicinity of the substrate binding region. These regions are marked by the letters 'X', 'Y', 'Z', and '$' (Fig. 5). Figure 6 shows the ribbon representation for two typical folds found for MMPs: one is for the human MMP-12 and another is MMP-19 (Fig. 6*A*, *B*, respectively). The three variable loops (at 10, 5, and 3 o'clock positions designated (in Fig. 5) as X, Z, and $, respectively) that could have contact with the bound substrates (vide infra) are shown in white in Fig. 6. The Y loop is located far from the substrate binding area on the

catalytic domain and is not especially highlighted in Fig. 6 (at 10 o'clock). The region designated by the letter X (at 10 o'clock in Fig. 6) is formed by the turn between the two antiparallel β-strands. These β-strands provide some of the binding ligands for the structural zinc and calcium binding sites. Our previous models for substrate binding in the active site of MMPs revealed that substrate can acquire an extended conformation (30). In such a binding mode, the unprimed portions [for convention on identification of substrates in protein–substrate complexes and their binding sites in proteases, consult Berger and Schechter (53)] of substrates would have contacts to the X loop. This loop is shorter in the human MMP-19 (Fig. 6*B*), frog XMMP, envelysins, plant MMPs, and nematode MMP than in all other MMPs. The shorter X loop makes the unprimed areas of the active site more open in these enzymes. The $ loop has contacts with the primed portion of the substrate (P_3' position) and is located at 3 o'clock in Fig. 6. The Z region (at 5 o'clock in Fig. 6) is an "Ω" loop, which forms the S_1' binding pocket; in the crystallized full-length porcine MMP-1, this loop has contacts through side chains and bridging water molecules to the hemopexin-like domain of the enzyme (54, 55). The length of this loop will control the size of the residue at P_1' position of the substrate. The composition of the loop will have an effect on substrate specificity (30). Furthermore, binding of protein substrates by some MMPs is influenced by interactions with the hemopexin-like domain (32). Nonetheless, the specificity of small synthetic substrates is triggered by their interactions solely with the active site in the catalytic domain and its surroundings. This is probably true even for gelatinases, since the catalytic domain of MMP-2 with excised gelatin binding domain is still active in hydrolysis of synthetic peptides (10, 56). Substrates interact with the loops designated X, $, and Z at unprimed and primed portions, and structural variability of these loops provide the diversity of such interactions.

CONCLUSION

The foregoing examined sequence similarities, sequence alignments, and structural aspects in arriving at an understanding of the important functions for the family of matrix metalloproteinases. Our analysis of the structures of MMPs in view of their evolutionary relationship follow the limitations of the primary sequence alignment and the prediction of the 3-dimensional structures, but present a distinctive way of looking at the multidomain structures like those of MMPs. The fact that alignment of the entire sequences and those for the catalytic domains and the hemopexin-like domains produced essentially the same numbers for the clusters (22–23 clusters) and that the composition for the clusters appear to be the same in each case is not coincidental. What this reveals to us is the likely scenario that domain assemblies occurred in an early stage of the diversification of these enzymes and that they progressed through the evolutionary process independent of one another, and perhaps in parallel to each other. This fact does not divorce itself from the obvious premise that at some primordial point in the evolution of these enzymes they must have existed as simple single-domain proteins that underwent gene fusions to generate the more complicated multidomain enzymes. This point is perhaps best underscored by the examples of the three plant MMPs and the sole enzyme from nematode. These clearly are modern variants of enzymes that did not undergo major structural elaboration in their development. However, our analysis also demonstrates that there are examples where evolution progressed in the reverse direction: a more complicated multidomain enzyme underwent truncation in its gene sequence to give rise to a less elaborated protein of fewer domains. An example of this type of retrograde process is matrilysin (MMP-7), which contains only the signal peptide, the propeptide, and the catalytic domain.

It is not clear how many more MMPs exist in nature, and our understanding of the actual functions of these enzymes is now at a rudimentary stage. As more sequences of MMPs become available, the analysis presented here should be updated and correlated with the new structural information that will be determined for these important enzymes. Nonetheless, the exercise presented here is the first step toward appreciation of the evolutionary processes that led to the diversification of these enzymes, with the attendant myriad of activities of central importance to both the physiology and pathology of living organisms. FJ

This work was supported by grants from the U.S. Army (to S.M.) and the Karmanos Cancer Institute (to S.M. and R.F.). R.F. is also supported by National Institutes of Health grant CA-61986. I.M. was a recipient of the Rumble and Heller predoctoral fellowships. We acknowledge Markku Kurkinen for providing us with the amino acid sequences for CMMP.

REFERENCES

1. Woodhouse, E. C., Chuaqui, R. F., and Liotta, L. A. (1997) General mechanisms of metastasis. *Cancer* 80 (Suppl. 8), 1529–1537
2. Chambers, A. F., and Matrisian, L. M. (1997) Changing views of the role of matrix metalloproteinases in metastasis. *J. Natl. Cancer Inst.* 89, 1260–1270
3. Coussens, L. M., and Werb, Z. (1996) Matrix metalloproteinases and the development of cancer. *Chem. Biol.* 3, 895–904
4. Yu, A. E., Hewitt, R. E., Connor, E. W., and Stetler-Stevenson, W. G. (1997) Matrix metalloproteinases. Novel targets for directed cancer therapy. *Drugs Aging* 11, 229–244
5. Kohn, E. C., and Liotta, L. A. (1995) Molecular insights into cancer invasion: strategies for prevention and intervention. *Cancer Res.* 55, 1856–1862

6. Bode, W., Reinemer, P., Huber, R., Kleine, T., Schnierer, S., and Tschesche, H. (1994) The X-ray crystal structure of the catalytic domain of human neutrophil collagenase inhibited by a substrate analogue reveals the essentials for catalysis and specificity. *EMBO J.* **13**, 1263–1269
7. Salowe, S. P., Marcy, A. I., Cuca, G. C., Smith, C. K., Kopka, I. E., Hagman, W. K., and Hermes, J. D. (1992) Characterization of zinc-binding sites in human stromelysin-1: stoichiometry of the catalytic domain and identification of a cysteine ligand in the proenzyme. *Biochemistry* **31**, 4535–4540
8. Borden, P., and Heller, R. A. (1997) Transcriptional control of matrix metalloproteinases and the tissue inhibitors of matrix metalloproteinases. *Crit. Rev. Eukaryot. Gen. Expression* **7**, 159–178
9. Gomis-Rüth, F.-X., Maskos, K., Betz, M., Bergner, A., Huber, R., Suzuki, K., Yoshida, N., Nagase, H., Brew, K., Bourenkov, G. P., Bartunik, H., and Bode, W. (1997) Mechanism of inhibition of the human matrix metalloproteinase stromelysin-1 by TIMP-1. *Nature (London)* **389**, 77–81
10. Murphy, G., Nguyen, Q., Cockett, M. I., Atkinson, S. J., Allan, J. A., Knight, C. G., Willenbrock, F., and Docherty, A. J. P. (1994) Assessment of the role of the fibronectin-like domain of gelatinase A by analysis of a deletion mutant. *J. Biol. Chem.* **269**, 6632–6636
11. Sato, H., Kinoshita, T., Takino, T., Nakayama, K., and Seiki, M. (1996) Activation of a recombinant membrane type 1-matrix metalloproteinase (MT1-MMP) by furin and its interaction with tissue inhibitor of metalloproteinases (TIMP)-2. *FEBS Lett.* **393**, 101–104
12. Puente, X. S., Pendás, A. M., Llano, E., Velasco, G., and López-Otín, C. (1996) Molecular cloning of a novel membrane-type matrix metalloproteinase from a human breast carcinoma. *Cancer Res.* **56**, 944–949
13. Santavicca, M., Noel, A., Angliker, H., Stoll, I., Segain, J. P., Anglard, P., Chretien, M., Seidah, N., and Basset, P. (1996) Characterization of structural determinants and molecular mechanisms involved in pro-stromelysin-3 activation by 4-aminophenylmercuric acetate and furin-type convertases. *Biochem. J.* **315**, 953–958
14. Wilhelm, S. M., Collier, I. E., Marmer, B. L., Eisen, A. Z., Grant, G. A., and Goldberg, G. I. (1989) SV40-transformed human lung fibroblasts secrete a 92-kDa type IV collagenase which is identical to that secreted by normal human macrophages. *J. Biol. Chem.* **264**, 17213–17221
15. Fridman, R., Bird, R. E., Hoyhtya, M., Oelkuct, M., Komarek, D., Liang, C.-M., Berman, M. L., Liotta, L. A., Stetler-Stevenson, W. G., and Fuerst, T. R. (1993) Expression of human recombinant 72 kDa gelatinase and tissue inhibitor of metalloproteinase-2 (TIMP-2): characterization of complex and free enzyme. *Biochem. J.* **289**, 411–416
16. Goldberg, G. I., Strongin, A., Collier, I. E., Genrich, L. T., and Marmer, B. L. (1992) Interaction of 92-kDa type IV collagenase with tissue inhibitor of metalloproteinases prevents dimerization, complex formation with interstitial collagenase, and activation of the proenzyme with stromelysin. *J. Biol. Chem.* **267**, 4583–4591
17. Butler, G. S., Butler, M. J., Atkinson, S. J., Will, H., Tamura, T., van Westrum, S. S., Crabbe, T., Clements, J., d'Ortho, M.-P., and Murphy, G. (1998) The TIMP2 membrane type 1 metalloproteinase 'receptor' regulates the concentration and efficient activation of progelatinase A. *J. Biol. Chem.* **273**, 871–880
18. Stöcker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Rüth, F.-X., McKay, D. B., and Bode, W. (1995) The metzincins—topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Protein Sci.* **4**, 823–840
19. Rawlings, N. D., and Barrett, A. J. (1995) Evolutionary families of metallopeptidases. *Methods Enzymol.* **248**, 183–228
20. Sang, Q. A., and Douglas, D. A. (1996) Computational sequence analysis of matrix metalloproteinases. *J. Protein Chem.* **15**, 137–160
21. Llano, E., Pendás, A. M., Knäuper, V., Sorsa, T., Salo, T., Salido, E., Murphy, G., Simmer, J. P., Bartlett, J. D., and López-Otín, C. (1997) Identification and structural and functional characterization of human enamelysin (MMP-20). *Biochemistry* **36**, 15101–15108
22. Bernot, A., Clepet, C., Dasilva, C., Devaud, C., Petit, J.-L., Caloustian, C., Cruaud, C., Samson, D., Pùlcini, F, Weissenbach, J., Heilig, R., Notanicola, C., Domingo, C., Rozenbaum, M., Benchetrit, E., Tpaloglu, R., Dewalle, M., Dross, C., Hadjari, P., Dupont, M., Demaille, J., Touitou, I., Smaoui, N., Nedelec, B., Méry, J.-P., Chaabouni, H., Delpech, M., and Grateau, G. (1997) A candidate gene for familial Mediterranean fever. *Nature Genet.* **17**, 25–31
23. Lepage, T., and Gache, C. (1990) Early expression of a collagenase-like hatching enzyme gene in the sea urchin embryo. *EMBO J.* **9**, 3003–3012
24. Namura, K., Shimuzu, T., Kinoh, H., Sendai, Y., Inomata, M., and Suzuki, N. (1997) Sea urchin hatching enzyme (Envelysin): cDNA cloning and deprivation of protein substrate specificity by autocatalytic degradation. *Biochemistry* **36**, 7225–7238
25. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hiller, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laisster, N., Latreille, P., Lightning, J., Lloyd, C., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., and Wohldman, P. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature (London)* **368**, 32–38
26. Lovejoy, B., Cleasby, A., Hassell, A. M., Longley, K., Luther, M. A., Weigl, D., McGeehan, G., McElroy, A. B., Drewry, D., Lambert, M. H., and Jordan, S. R. (1994) Structure of the catalytic domain of fibroblast collagenase complexed with an inhibitor. *Science* **263**, 375–377
27. Stams, T., Spurlino, J. C., Smith, D. L., Wahl, R. C., Ho, T. F., Qoronfleh, M. W., Banks, T. M., and Rubin, B. (1994) Structure of human neutrophil collagenase reveals large S_1' specificity pocket. *Nature Struct. Biol.* **1**, 119–123
28. Gooley, P. R., O'Connell, F., Marcy, A. I., Cuca, G. C., Salowe, S. P., Bush, B. L., Hermes, J. D., Esser, C. K., Hagmann, W. K., Springer, J. P., and Johnson, B. A. (1994) The NMR structure of the inhibited catalytic domain of human stromelysin-1. *Nature Struct. Biol.* **1**, 111–118
29. Browner, M. F., Smith, W. W., and Castelhano, A. L. (1995) Matrilysin-inhibitor complexes: common themes among metalloproteases. *Biochemistry* **34**, 6602–6610
30. Massova, I., Fridman, R., and Mobashery, S. (1997) Structural insight into the catalytic domains of human matrix metalloprotease-2 and human matrix metalloprotease-9: implications for substrate specificities. *J. Mol. Mod.* **3**, 17–30
31. Massova, I., Kotra, L. P., and Mobashery, S. (1998) Structural insight into the binding motifs for the calcium ion and the non-catalytic zinc in matrix metalloproteases. *Bioorg. Med. Chem. Lett.* **8**, 853–858
32. Murphy, G., and Knäuper, V. (1997) Relating matrix metalloproteinase structure to function: why the 'hemopexin' domain? *Matrix Biol.* **15**, 511–518
33. Murphy, G., Allan, J. A., Willenbrock, F., Crockett, M. I., O'Connell, J. P., and Docherty, J. P. (1992) The role of the C-terminal domain in collagenase and stromelysin specificity. *J. Biol. Chem.* **267**, 9616–9618
34. Murphy, G. J., Murphy, G., and Reynolds, J. J. (1991) The origin of matrix metalloproteinases and their familial relationships. *FEBS Lett.* **289**, 4–7
35. De Souza, S. J., and Brentani, R. (1993) Sequence homology between a bacterial metalloproteinase and eukaryotic matrix metalloproteinases. *J. Mol. Evol.* **36**, 596–598
36. McGeehan, G., Burkhart, W., Anderegg, R., Becherer, D., Gillikin, J. W., and Graham, J. S. (1992) Sequencing and characterization of the soybean leaf metalloproteinase. *Plant Physiol.* **99**, 1179–1183
37. Franco, A. A., Mundy, L. M., Trucksis, M., Wu, S., Kaper, J. B., and Sears, C. L. (1997) Cloning and characterization of the *Bacteroides fragilis* metalloprotease toxin gene. *Infect. Immunol.* **65**, 1007–1013
38. Yang, M., Murray, M. T., and Kurkinen, M. (1997) A novel metalloproteinase gene (XMMP) encoding vitronectin-like motifs is transiently expressed in *Xenopus laevis* early embryo development. *J. Biol. Chem.* **272**, 13527–13533

39. Pei, D., and Weiss, S. J. (1995) Furin-dependent intracellular activation of the human stromelysin-3 zymogen. *Nature (London)* **375**, 244–247
40. Van Wart, H. E., and Birkedal-Hansen, H. (1990) The cysteine switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family. *Proc. Natl. Acad. Sci. USA* **87**, 5578–5582
41. Becker, J. W., Marcy, A. I., Rokosz, L. L., Axel, M. G., Burbaum, J. J., Fitzgerald, P. M. D., Cameron, P. M., Esser, C. K., Hagmann, W. K., Hermes, J. D., and Springer, J. P. (1995) Stromelysin-1: three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme. *Protein Sci.* **4**, 1966–1976
42. Bode, W., Grams, F., Reinemer, P., Gomis-Rüth, F. X., Baumann, U., McKay, D. B., and Stocker, W. (1996) The metzincin-superfamily of zinc-peptidases. *Adv. Exp. Med. Biol.* **389**, 1–11
43. Olson, M. W., Gervasi, D. C., Mobashery, S., and Fridman, R. (1997) Kinetic analysis of the binding of human matrix metalloproteinase-2 and -9 to tissue inhibitor of metalloproteinase (TIMP)-1 and TIMP-2. *J. Biol. Chem.* **272**, 29975–29983
44. Strongin, A. Y., Collier, I. E., Bannikov, G., Marmer, B. L., Grant, G. A., and Goldberg, G. I. (1995) Mechanism of cell surface activation of 72-kDa type IV collagenase. *J. Biol. Chem.* **270**, 5331–5338
45. Gohlke, U., Gomis-Rüth, F. X., Crabbe, T., Murphy, G., Docherty, A. J., and Bode, W. (1996) The C-terminal (hemopexin-like) domain structure of human gelatinase A (MMP-2): structural implications for its function. *FEBS Lett.* **378**, 126–130
46. Brooks, P. C., Stromblad, S., Sanders, L. C., von Schalscha, T. L., Aimes, R. T., Stetler-Stevenson, W. G., Quigley, J. P., and Cheresh, D. A. (1996) Localization of matrix metalloproteinase-2 to the surface of invasive cells by interaction with integrin $\alpha_v\beta_3$. *Cell* **85**, 683–693
47. Brooks, P. C., Silletti, S., von Schalscha, T. L., Friedlander, M., and Cheresh, D. A. (1998) Disruption of angiogenesis by PEX, a noncatalytic metalloproteinase fragment with integrin binding activity. *Cell* **92**, 391–400
48. Namura, K., Tanaka, H., Kikkawa, Y., Yamaguchi, M., and Suzuki, N. (1991) The specificity of sea urchin hatching enzyme (envelysin) places it in the mammalian matrix metalloproteinases. *Biochemistry* **30**, 6115–6123
49. Bartlett, J. D., Simmer, J. P., Xue, J., Margolis, H. C., and Moreno, E. C. (1996) Molecular cloning and mRNA tissue distribution of a novel matrix metalloproteinase isolated from porcine enamel organ. *Gene* **183**, 123–128
50. Stolow, M. A., Bauzon, D. D., Li, J., Sedgwick, T., Liang, V. C.-T., Sang, Q. A., and Shi, Y.-B. (1996) Identification and characterization of a novel collagenase in *Xenopus laevis:* possible roles during frog development. Mol. Biol. Cell. **7**, 1471–1483
51. Miyazaki, K., Uchiyama, K., Imokawa, Y., and Yoshizato, K. (1996) Cloning and characterization of cDNAs for matrix metalloproteinases of regenerating newt limbs. *Proc. Natl. Acad. Sci. USA* **93**, 6819–6824
52. Noel, A., Santavicca, M., Stoll, I., L'Hoir, C., Staub, A., Murphy, G., Rio, M. C., and Basset, P. (1995) Identification of structural determinants controlling human and mouse stromelysin-3 proteolytic activities. *J. Biol. Chem.* **270**, 22866–22872
53. Berger, A., and Schechter, I. (1970) Mapping the active site of papain with the aid of peptide substrates and inhibitors. *Phil. Trans. R. Soc. London B Biol. Sci.* **257**, 249–264
54. Li, J., Brick, P., O'Hare, M. C., Skarzyrski, T., Lioyd, L. F., Curry, V. A., Clark, I. M., Bigg, H. F., Hazelman, B. L., Cawston, T. E., and Blow, D. M. (1995) Structure of full-length porcine synovial collagenase reveals a C-terminal domain containing a calcium-linked, four-bladed β-propeller. *Structure* **3**, 541–548
55. Bode, W. (1995) A helping hand for collagenases: the hemopexin-like domain. *Structure* **3**, 527–530
56. Ye, Q. Z., Johnson, L. L., Yu, A. E., and Hupe, D. (1995) Reconstructed 19 kDa catalytic domain of gelatinase A is an active proteinase. *Biochemistry* **34**, 4702–4708
57. Kräulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950
58. Merritt, E. A., and Bacon, D. J. (1997) Raster3D photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524

Pergamon

Bioorganic & Medicinal Chemistry Letters 8 (1998) 853–858

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

STRUCTURAL INSIGHT INTO THE BINDING MOTIFS FOR THE CALCIUM ION AND THE NON-CATALYTIC ZINC IN MATRIX METALLOPROTEASES

Irina Massova, Lakshmi P. Kotra, and Shahriar Mobashery*

Department of Chemistry, Wayne State University, Detroit, MI 48202-3489, U.S.A.

Received 19 January 1998; accepted 23 February 1998

Abstract: The binding motifs for the structural zinc and calcium ions in matrix metalloproteases (MMPs) were investigated by analyzing the three-dimensional structural models of 23 representative MMPs.

Matrix metalloproteinases (MMPs) constitute a major group of proteases that carry out a myriad of physiological and pathological functions. These functions include embryogenesis, angiogenesis, wound healing, inflammation, arthritis, and cancer metastasis.[1,2] These enzymes are zinc-dependent endopeptidases known for their ability to cleave extra-cellular matrix constituents, as well as non-matrix proteins. MMPs have developed into a unique group of zinc-dependent proteinases characterized by the incorporation of various protein domains in their structures, which mediate interactions with substrates and inhibitors.[3] The implication of MMPs in cancer metastasis and angiogenesis has raised considerable interest in the MMP family since they represent attractive targets for development of novel anticancer drugs. Therefore, understanding of the structure and function of these key enzymes would have significant implications toward these efforts.

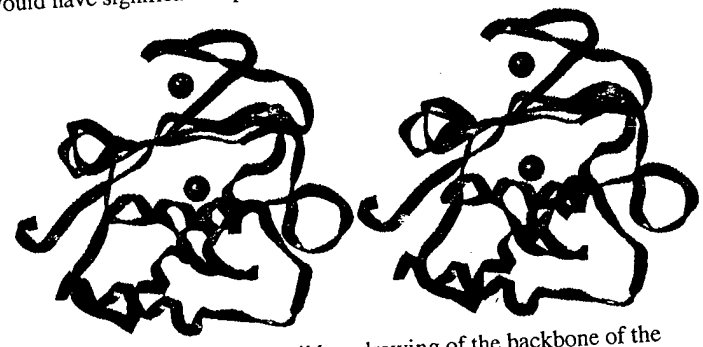


Figure 1. Stereo view of the ribbon drawing of the backbone of the catalytic domain of MMP-19. The red sphere represents the catalytic zinc ion, the green sphere represents the structural zinc ion and the orange sphere represents the calcium ion.

A total of 64 MMPs have been sequenced to date, of which 15 are from human. These human enzymes have counterparts in other vertebrates. Furthermore, MMPs have even been identified in invertebrates[4-6] and three have recently been sequenced from plant sources.[7] MMPs, in general, possess a propeptide domain (N-terminal), a catalytic domain and a hemopexin-like domain (C-terminal; except for MMP-7 which lacks it).[8] MMP-2 and -9 further possess a fibronectin-like domain and membrane-type MMPs (MT-MMPs) have acquired a transmembrane domain. In the catalytic domain, there are two zinc ions and at least one calcium ion coordinated to various residues (Figure 1). One of the two zinc ions is present in the active site, and is intimately involved in

0960-894X/98/$19.00
PII: S0960-894X(98)00128-0

the catalytic processes of these enzymes. The second zinc ion (structural zinc) and the calcium ion are present in the catalytic domain in proximity to the catalytic zinc (approximately 12 Å away). The catalytic zinc ion is essential for the proteolytic activity of MMPs and the three histidine residues that coordinate with the catalytic zinc are conserved among MMPs. However, there is little known about the roles of the second zinc ion and the calcium ion within the catalytic domain. It was noted that the MMPs have high affinities for the structural zinc and the calcium ions.[9,10] It has been suggested that these metal ions keep the structural elements together in the catalytic domains of the MMPs, contributing to its stability.[11-13] We have compared amino-acid sequences of 64 MMPs from various sources. Whereas the catalytic zinc and its coordinating residues are well conserved, structural zinc and the calcium ions show differences in their corresponding binding motifs. In an effort to understand the coordination patterns of the structural zinc and calcium ions in the catalytic domains of various MMPs, we have modeled the three-dimensional structures of several representative members of the MMPs. A clear insight into these coordination sites is important due to the fact that the structural zinc-binding motif as well as the calcium-binding motif are unique to the matrix metalloproteases and are not present in other metzincin proteases like astacin and adamalysin.[11] This comprehensive analysis of the three-dimensional models of the structural zinc and calcium ion binding regions has been performed here for the first time to expand our knowledge of this important family of enzymes.

Amino-acid sequences of MMPs were obtained from the GenBank, TREMBL, and SwissProt data banks. We utilized a total of 64 MMP sequences from various sources for the multiple-sequence alignment and then, a separate analysis was run using the representative 23 MMP sequences (*vide infra*). The human enzymes were selected where possible. Four human MMPs, fibroblast (MMP-1, 1cgl)[12] and neutrophil (MMP-8, 1mnc)[14] collagenases, matrilysin (MMP-7, 1mmq),[15] and stromelysin-1 (MMP-3, 1slm)[16] have recently been crystallized. The 17 modeled MMPs are the human MMP-10, MMP-11, MMP-12, MMP-13, MT1-MMP (MMP-14), MT2-MMP (MMP-15), MT3-MMP (MMP-16), MT4-MMP (MMP-17), MMP-19 (same as MMP-18), pig enamelysin, sea urchin envelysin (SwissProt accession number P22757), stromelysin-like MMP from newt (GenBank accession number D82053), collagenase-4 from frog, nematode MMP (GenBank accession number U00038), chicken CMMP, frog XMMP, and MMP from mouse-ear cress (TREMBL accession number O04529). The multiple-sequence alignments were performed using the program PileUp from the Wisconsin package version 9. We used the existing X-ray crystal structural information to predict the three-dimensional structures for the catalytic domains of 17 homologous metalloproteases using the program COMPOSER.[17] Water molecules that occupy the average positions in the crystallized MMPs and the two zinc and one calcium ions were added to the folded structures, and the entire complexes were allowed to undergo 10000 cycles of energy-minimization according to the methodology reported by us earlier.[18] We have reported recently on such computational structures for the catalytic domains of gelatinases A (MMP-2) and B (MMP-9).[18] After modeling the catalytic domains of the 17 MMPs, we utilized this information in conjunction with the four X-ray crystal structures of MMPs and previously modeled MMP-2 and MMP-9 structures to compare the structural zinc-binding and calcium-binding regions.

Figure 2 shows the multiple-sequence alignment for the structural zinc-binding and calcium-binding sites of the catalytic domains of the representative 23 MMPs. These metal-binding sites are formed by β-strands and turns. Analysis of the second zinc-binding site revealed that the MMPs developed at least four different motifs for

binding to this structural zinc ion. For the majority of MMPs (except for MMP-11, MMP-17, CMMP and XMMP), this site consists of the side chains of an aspartic acid and three histidine residues; which are consistent with the existing crystal structures. Two of the coordinating histidine imidazoles are bound as His_δ and one as His_ε (Figure 2; the three-dimensional structure of this site is shown in Figure 3A). The signature for this site in the 60 MMPs is the sequence(s) ***H*$_\delta$**[GN]-***D***-X(2)-[PAS]-**F**-***D***-[GA]-X(4)-[LIRV]-[AG]-***H*$_\delta$**[AV]-[**FYS**]-P-X(5,7,9)-***H*$_\varepsilon$**-[FL]-***D***-X(2)-***E***-X-**W**, where residues that provide side chains for coordination to zinc and calcium are shown in bold italics, X (5,7,9) indicates 5, 7, or 9 variable residues in between the flanking sites (H_δ is a histidine protonated at $N_{\delta 1}$ and H_ε is a histidine protonated at $N_{\varepsilon 2}$). The first phenylalanine and the second phenylalanine/tyrosine (i.e., the FYS sequence) in this pattern are brought close to the coordinated histidines by the enzyme fold. This creates an increased hydrophobic environment, which presumably would enhance the binding affinity for the metal ions. The human MMP-11 has an aspartic acid instead of one of the His_δ. This is in contrast to the MMP-11 from rabbit and mouse, which still have a histidine at this position (data not shown). Therefore, we conclude that MMP-11 (compared to those from human, rabbit and mouse) does not have a unique motif for binding of the structural zinc.

```
MMP-1 human       DIMISFVRGDHRDNSP-FDG PGGNLAHAFQPGPGIGGD-AHFDEDERWTNN---------
MMP-2 human       DIMINFGRWEHGDGYP-FDG KDGLLAHAFAPGTGVGGD-SHFDDDELWTLGEGQVVRVKY
MMP-3 human       DIMISFAVREHGDFYP-FDG PGNVLAHAYAPGPGINGD-AHFDDDEQWTKD---------
MMP-7 human       DIMIGFARGAHGDSYP-FDG PGNTLAHAFAPGTGLGGD-AHFDEDERWTDG---------
MMP-8 human       DINIAFYQRDHGDNSP-FDG PNGILAHAFQPGQGIGGD-AHFDAEETWTNT---------
MMP-9 human       DIVIQFGVAEHGDGYP-FDG KDGLLAHAFPPGPGIQGD-AHFDDDELWSLGKGVVVPTRF
MMP-10 human      DIMISFAVKEHGDFYS-FDG PGHSLAHAYPPGPGLYGD-IHFDDDEKWTED---------
MMP-11 human      DIMIDFARYWDGDDLP-FDG PGGILAHAFFPKTHREGD-VHFDYDETWTIGDDQ------
MMP-12 human      DILVVFARGAHGDFHA-FDG KGGILAHAFGPGSGIGGD-AHFDEDEFWTTH---------
MMP-13 human      DIMISFGIKEHGDFYP-FDG PSGLLAHAFPPGPNYGGD-AHFDDDETWTSS---------
MT1-MMP human     DIMIFFAEGFHGDSTP-FDG EGGFLAHAYFPGPNIGGD-THFDSAEPWTVRNEDL-----
MT2-MMP human     DIMVLFASGFHGDSSP-FDG TGGFLAHAYFPGPGLGGD-THFDADEPWTFSSTDL-----
MT3-MMP human     DIPIIFASGFHGDSSP-FDG EGGFLAHAYFPGPGIGGD-THFDSDEPWTLGNPNH-----
MT4-MMP human     DIQIDFSKADHNDGYP-FDA -RRHRAHAFFPGHHHTAGYTHFNDDEAWTFRSSDA-----
MMP19 human       DIRLSFHGRQSSYCSNTFDG PGRVLAHADIPELG----SVHFDEDEFWTEGTY-------
Collagenase-4 frog DIEISFTAGDHKDNSP-FDG SGGILAHAFQPGNGIGGD-AHFDEDETWTKT---------
MMP newt          DIQISFGAREHGDFNP-FDG PYGTLAHAFAPGTGIGGD-AHFDEDEKWSKV---------
CMMP chicken      DIMVAFGTKAHGHCPRYFDG PLGVLAHAFPPGSGFGGD-VHFDEDEDWTMG---------
Enamelysin pig    DIMISFETGDHGDSYP-FDG PRGTLAHAFAPGEGLGGD-THFDNAEKWTMG---------
Envelysin         DIRIKFGSYDHGDGIS-FDG RGGVLAHAFLPRNG----DAHFDDSETWTEGTR--------
MMP nematode      DIYIAFEKGEHSDGFP-FDG QDGVVAHAFYPRDG----RLHFDAEEQWSLNSV--------
XMMP frog         DIKLGFGRGRHLGCSRAFDG SGQEFAHAWFLGD------IHFDDDEHFTAPS---------
MMP cress         DITIGFYTGDHGDGEP-FDG VLGTLAHAFSPPSG----KFHLDADENWVVSGDL------
```

Figure 2. Multiple-sequence alignment of the structural zinc and calcium binding regions in the catalytic domains of 23 representative MMPs. The structural zinc binding residues in majority of MMPs are given in blue - three histidines and one aspartic acid; the structural zinc binding domain in MMP-19, CMMP, and XMMP are in green. The residues that coordinate to the calcium ion are in pink; "underlined" amino acids chelate to the calcium ion via their backbone carbonyl moiety; the variation which was observed in the calcium binding site of MMP-17 (MT4-MMP) is given in cyan.

The remaining MMPs, those of chicken CMMP, the frog XMMP, and the human MMP-19 (Figure 2), each possess unique motifs for binding to the structural zinc ion. The chicken CMMP has a histidine where the

majority of MMPs have an aspartic acid coordinated to the second zinc ion. Consequently, the structural zinc ion in this enzyme is coordinated by four histidine residues (His-171, His-173, His-187 and His-200; Figure 3B). Such coordination would force the cysteine adjacent to the fourth coordinated histidine (Cys-174) to point to the outside milieu. This cysteine may provide a site for protein dimerization, as it is entirely exposed and available. Such active dimer formation has been reported for MMP-9, for example.[19] We also investigated the possibility for the existence of another binding mode for the second zinc ion, where the three original histidines and the cysteine are coordinated to the zinc ion (Figure 3C). Such structural alternative is a distinct possibility, and it would still be a novel motif for coordination to the structural zinc. The frog XMMP has a cysteine residue at the same position as in the sequence of CMMP. However, it is missing the aspartic acid at the usual location typical for most MMPs for coordination to the second zinc ion. This enzyme has a glycine in the place of aspartic acid, which obviously cannot provide metal coordination (Figure 2). Therefore, for XMMP the only possibility is to have three histidines (His-270, His-286 and His-294) and the cysteine (Cys-273) side chains to coordinate to the second zinc ion (Figure 3D).

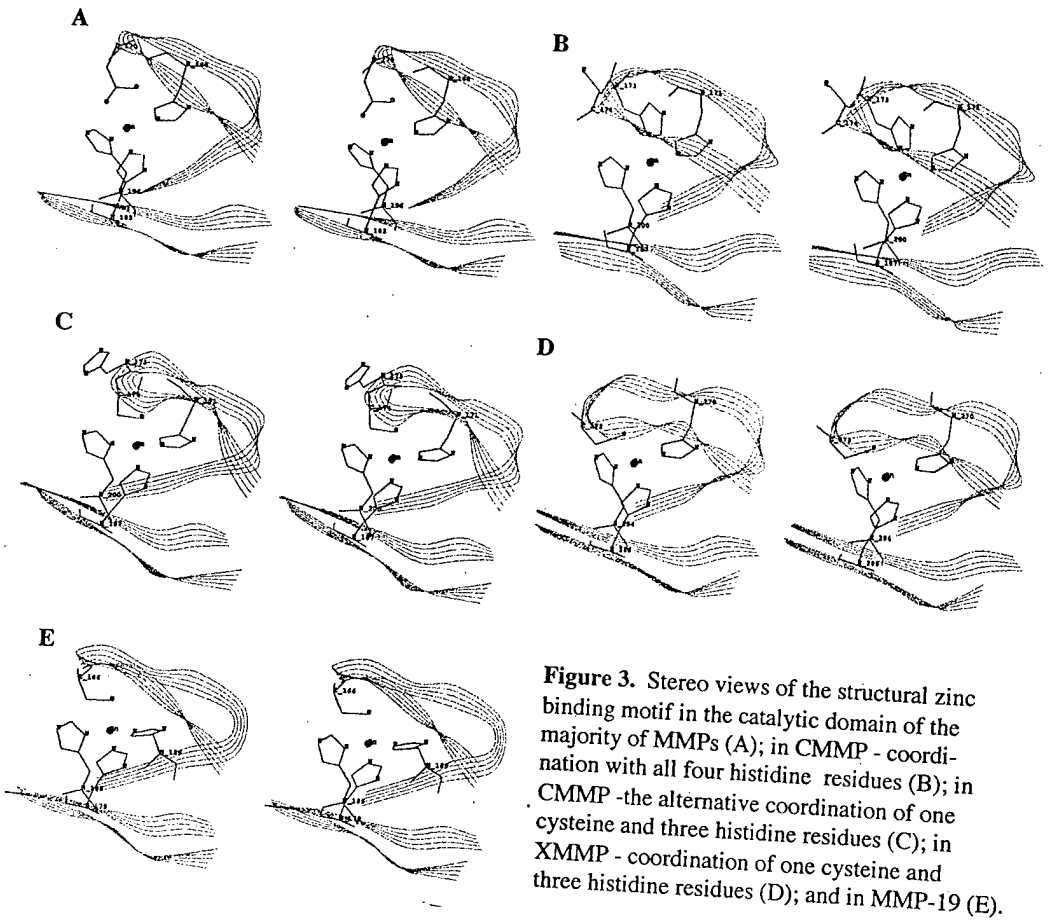


Figure 3. Stereo views of the structural zinc binding motif in the catalytic domain of the majority of MMPs (A); in CMMP - coordination with all four histidine residues (B); in CMMP -the alternative coordination of one cysteine and three histidine residues (C); in XMMP - coordination of one cysteine and three histidine residues (D); and in MMP-19 (E).

Analysis of the structure for MMP-19 showed that this enzyme has two histidines for metal coordination at the usual places and one cysteine at the corresponding position already discussed in the sequences for CMMP and XMMP. However, MMP-19 possesses a non-coordinating serine (Ser-163) at the position where the majority of MMPs have the third histidine. Interestingly, the third coordinating histidine (His-159, Figure 3E) in this enzyme is provided at an entirely distinct position, four residues to the N-terminus. Whereas the position of the amino acid is different and indeed on a different β-strand, it provided the side chain for the fourth coordination site in a perfectly acceptable orientation in space (for example, compare the positions of His-159 and His-270 in Figures 3E and 3D, respectively).

Binding of the calcium ion brings six specific elements of MMP catalytic domain for coordination in an octahedral fashion. Three of these elements are the side chains of two aspartic acids and one glutamic acid, which are conserved in all 23 MMPs except in MT4-MMP (*vide infra*, Figure 2). The remaining three calcium ligands are provided by the backbone carbonyl oxygens of three residues within a turn made up of five amino acids (Figure 2). A typical calcium-binding motif is shown in Figure 4A. The only exception to this general picture would appear to be MT4-MMP (MMP-17), which has undergone one amino acid deletion in this turn (i.e., it has a four-amino acid turn). In human MMP-17 (MT4-MMP), an asparagine residue is observed instead of the second aspartic acid (Figure 2). The coordination to the calcium ion for MT4-MMP (MMP-17) is shown in Figure 4B.

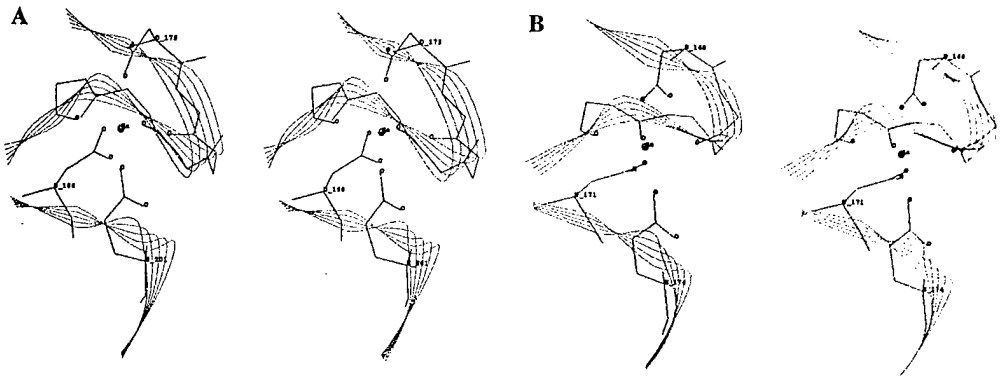


Figure 4. Stereo views of the calcium binding motif in majority of MMPs, except for MMP-17 (A); that for MMP-17 (B).

In conclusion, there are at least four different motifs for the binding to the structural zinc ion that are discernible from the three-dimensional structure analysis. However, the calcium binding motif is more strictly conserved. These variations in the metal-binding motifs preserve the topology of the structural zinc-binding site, as well as the calcium-binding site, and hence the general fold for the catalytic domain is highly preserved. The second zinc binding site is present in all MMPs and is important for the activity of these enzymes. Hence, the structural variations that are noted here in the coordination to the structural zinc ion may be indicative of the different outcomes for selection of novel enzymic activities. Furthermore, dendrogram analysis of the catalytic

domain sequence alignment for MMPs indicates that these variations in the zinc-binding motif came about as a consequence of independent evolutionary processes, unrelated to one another.[20] Since these motifs are absent in other members of the metzincin family of enzymes, we venture to say that the existence of these metal-binding sites in MMPs must have arisen in response to specific needs unique to MMPs with reference to their substrate specificities. The information provided herein is intended to stimulate interest in exploring the role of these motifs in MMPs geared toward understanding the substrate specificities for these enzymes, which are not understood to any appreciable degree at the present. Our understanding of the actual functions of the MMPs is at a rudimentary stage at the present and as more sequences of MMPs become available, the analysis presented here should be updated and correlated with the new structural information to shed light on the various properties of these important enzymes.

Acknowledgments. This work was supported by a grant from the US Army (DAMD17-97-1-7174). IM was a recipient of the Rumble and Heller predoctoral fellowships. We acknowledge Markku Kurkinen for providing us with the amino-acid sequence for CMMP.

References.

1. Chen, W. -T. *Curr. Opinion Cell Biol.* **1992**, *4*, 802.
2. Hagmann, W. K.; Lark, M. W.; Becker, J. W. In *Ann. Rev. Med. Chem.*; Bristol, J. A., Ed.; Academic Press, 1996; Vol. 31, pp 231-240.
3. Baramova, E.; Foidart, J. -M. *Cell Biol. Int.* **1995**, *19*, 239.
4. Lepage, T.; Gache, C. *EMBO J.* **1990**, *9*, 3003.
5. Namura, K.; Shimuzu, T.; Kinoh, H.; Sendai, Y.; Inomata, M.; Suzuki, N. *Biochemistry* **1997**, *36*, 7225.
6. Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, J.; Connell, M.; Copsey, T.; Cooper, J.; Coulson, A.; Craxton, M.; Dear, S.; Du, Z.; Durbin, R.; Favello, A.; Fraser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hiller, L.; Jier, M.; Johnston, L.; Jones, M.; Kershaw, J.; Kirsten, J.; Laisster, N.; Latreille, P.; Lightning, J.; Lloyd, C.; Mortimore, B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.; Shownkeen, R.; Sims, M.; Smaldon, N.; Smith, A.; Smith, M.; Sonnhammer, E.; Staden, R.; Sulston, J.; Thierry-Mieg, J.; Thomas, K.; Vaudin, M.; Vaughan, K.; Waterston, R.; Watson, A.; Weinstock, L.; Wilkinson-Sproat, J.; Wohldman, P. *Nature* **1994**, *368*, 32.
7. McGeehan, G.; Burkhart, W.; Anderegg, R.; Becherer, D.; Gillikin, J. W.; Graham, J. S. *Plant Physiol.* **1992**, *99*, 1179.
8. Murphy, G.; Knäuper, V. *Matrix Biol.* **1997**, *15*, 511.
9. Bode, W.; Reinemer, P.; Huber, R.; Kleine, T.; Schnierer, S.; Tschesche, H. *EMBO J.* **1994**, *13*, 1263.
10. Salowe, S. P.; Marcy, A. I.; Cuca, G. C.; Smith, C. K.; Kopka, I. E.; Hagman, W. K.; Hermes, J. D. *Biochemistry* **1993**, *31*, 4535.
11. Willenbrock, F.; Murphy, G.; Phillips, I. R.; Brocklehurst, K. *FEBS Lett.* **1995**, *358*, 189.
12. Lovejoy, B.; Cleasby, A.; Hassell, A. M.; Longley, K.; Luther, M. A.; Weigl, D.; McGeehan, G.; McElroy, A. B.; Drewry, D.; Lambert, M. H.; Jordan, S. R. *Science* **1994**, *263*, 375.
13. Wetmore, D. R.; Hardman, K. D. *Biochemistry* **1996**, *35*, 6549.
14. Browner, M. F.; Smith, W. W.; Castelhano, A. L. *Biochemistry* **1995**, *34*, 6602.
15. Gooley, P. R.; O'Connell, F.; Marcy, A. I.; Cuca, G. C.; Salowe, S. P.; Bush, B. L.; Hermes, J. D.; Esser, C. K.; Hagmann, W. K.; Springer, J. P.; Johnson, B. A. *Nature Struct. Biol.* **1994**, *1*, 111.
16. Stams, T.; Spurlino, J. C.; Smith, D. L.; Wahl, R. C.; Ho, T. F.; Qoronfleh, M. W.; Banks, T. M.; Rubin, B. *Nature Struct. Biol.* **1994**, *1*, 119.
17. Coordinates for the three-dimensional structures of the 17 MMPs that were modeled will be made available from our group web site, http://sun2.chem.wayne.edu/~somgroup upon acceptance of this manuscript.
18. Massova, I.; Fridman, R.; Mobashery, S. *J. Mol. Model.* **1997**, *3*, 17.
19. Clegg, P. D.; Burke, R. M.; Coughlan, A. R.; Riggs, C. M.; Carter, S. D. *Equine Vet. J.* **1997**, *29*, 343.
20. Massova, I.; Kotra, L. P.; Fridman, R.; Mobashery, S. *FASEB J.* submitted.